Tips and tricks for manual model building of atomic models into cryoEM maps. Oliver Clarke

An atomic model is a compact interpretation of the density map in light of prior knowledge (both specific and general).

- Aim is to build a model that is consistent with **both** the density map and everything we independently know about the structure/composition of the macromolecule of interest, both specifically and in terms of our general knowledge of protein structure and chemistry.
- At medium resolution (3-5 Å), this still requires manual building. Even the best autobuilt model still requires a lot of manual inspection and correction in most cases. (generates many fragments which need inspection, correction, merging)
- Tradeoff between available prior knowledge and required resolution for atomic modelling at the extremes, if a complete crystal structure is already available, 10Å data may be sufficient, while if no sequence/composition data is available even 3Å may not suffice.

Prior knowledge

- Protein sequence and derived info (secondary structure predictions, covariation/conservation, patterns of large/aromatic residues), disorder & contact prediction
- Crystal structures (+ homology models)
- Knowledge of protein structure, folding, chemistry, geometry.

Density map

- Resolution (+ local resolution, + map modification/sharpening)
- Patterns of large/small/absent sidechains
- Sharpening and density modification
- Conformational/compositional heterogeneity



- If possible, unique model that agrees with both density map and priors
- Otherwise (and per region), specify ambiguity (w/UNK residues and numbering or Ca only model)
- Validation not just (or even mostly) about overfitting.
- Identify, analyse, fix errors.
- Direction and register of sequence fit.
- Ligand identification/assignment.
- No model is or ever will be perfect. That's okay.

Before you start – make sure your maps are appropriately sharpened and low pass filtered! (and consider whether building is justified or whether further improvement of the reconstruction is required first)

- Often it is helpful to build using multiple maps. Assuming 3-3.5Å global res, I would suggest using a map filtered to the global resolution, one filtered to the best local resolution, and one filtered to ~4-4.5 Å (to better visualize connectivity).
- Try both simple B-factor sharpening and the approach used by *phenix.auto_sharpen*, which incorporates anisotropy removal. CisTEM *sharpen_map* also seems to give very good results in some cases.
- Also, if your map doesn't "look like" 4 Å, trust your eyes! If it is nominally 4Å and there are no sidechains visible, or your helices look "stretched", assess orientation bias (3D-FSC server: <u>https://3dfsc.salk.edu</u>), local resolution variation, and double check sharpening and masking parameters (are you *sure* you're looking at the sharpened map? Is the mask used for FSC calculation sensible?)

- Start by identifying boundaries of conserved domains (NCBI CDD: <u>https://www.ncbi.nlm.nih.gov/Structure/cdd/</u>; DELTA-BLAST also performs CD-search by default)
- Then identify suitable structural templates for building known domains: FUGUE, SPARKS-X, PHYRE2, MUSTER.
- Secondary structure, TM & disorder prediction (XtalPRED for overall summary; specific tools such as SPOT-DISORDER, SPIDER3 for best accuracy).
- Contact prediction from evolutionary couplings: EVFOLD & GREMLIN.
- Conservation analysis: Use favorite MSA algorithm (MUSCLE & CLUSTAL-OMEGA work well; TM-COFFEE, PRALINE-TM useful for membrane proteins) to create a sequence alignment of your protein with a few orthologs; gaps & insertions most commonly occur in loops/disordered regions. Useful as a guide during building.

- Start by identifying boundaries of conserved domains (NCBI CDD: <u>https://www.ncbi.nlm.nih.gov/Structure/cdd/</u>; DELTA-BLAST also performs CD-search by default)
- Then identify suitable structural templates for building known domains: FUGUE, SPARKS-X, PHYRE2, etc.
- Secondary structure, TM & disorder prediction (XtalPRED for overall summary; specific tools such as SPOT-DISORDER, SPIDER3 for best accuracy).
- Contact prediction from evolutionary couplings: EVFOLD & GREMLIN.
- Conservation analysis: Use favorite MSA algorithm (MUSCLE & CLUSTAL-OMEGA work well) to create a sequence alignment of your protein with a few orthologs; gaps & insertions most commonly occur in loops/disordered regions. Useful as a guide during building.

- Start by identifying boundaries of conserved domains (NCBI CDD: <u>https://www.ncbi.nlm.nih.gov/Structure/cdd/</u>; DELTA-BLAST also performs CD-search by default)
- Then identify suitable structural templates for building known domains: FUGUE, SPARKS-X, PHYRE2, etc.
- Secondary structure, TM & disorder prediction (XtalPRED for overall summary; specific tools such as SPOT-DISORDER, SPIDER3 for best accuracy).
- Contact prediction from evolutionary couplings: EVFOLD & GREMLIN.
- Conservation analysis: Use favorite MSA algorithm (MUSCLE & CLUSTAL-OMEGA work well; TM-COFFEE, PRALINE-TM useful for membrane proteins) to create a sequence alignment of your protein with a few orthologs; gaps & insertions most commonly occur in loops/disordered regions. Useful as a guide during building.

CDD provides a guide to domain level architecture, including sequence alignments & representative structures.

			Domain	S WA SH3		SH2		27373965	2	
IE SEARCH GUID	NewSe	arch Structure Hor	me 3	D Macromolecular S	tructures	Cons	erved Do	omains	Pubchem	BioSyste
Conserved	domains yanodine rec	on [sp P21817] ceptor 1 OS=Homo sag	piens OX=960	06 GN=RYR1 PE=1	SV=3			Vie	W Standard Res	sults ᅌ
Protein Class PRY1_RyR and I	ification RR_TM4-6 do	main-containing proteir	n (domain archi	tecture ID 11696388)						
rotein containing o	domains RYDR	_ITPR, SPRY1_RyR, SPRY2	2_RyR, and RR_	TM4-6						
Graphical sui	mmary 📋	Zoom to residue leve	show extra	options »						
	t	1000	20	••	3000		4000		5438	
uery seq.			2001							
on-specific	In	SPRY SPR S	SPRY						Ion_t	
its	THE REAL PROPERTY AND ADDRESS OF									
Superfamilies	I I R1	IDR_ SPRY Ry Ry SPR S	PRY	RYDR_I Ry	Ry		RIH	RR_TH4-	-6 Ion_t	
			Search for sir	nilar domain architectu	es 🛛 F	Refine search	2			
List of domai	n hits		Search for sir	nilar domain architectur	es 7 F	Refine search	2		Interval	E-val
List of domai	Accession	RIH domain: The RIH (Ry	Search for sir	nilar domain architectur Descripti nology) domain is an ex	es 7 F	Refine search	Des		Interval 443-636	E-val
List of domai	n hits Accession pfam01365 pfam01365	RIH domain; The RIH (Ry RIH domain; The RIH (Ry	Search for sir	nilar domain architectur Descripti nology) domain is an ex nology) domain is an ex	es 2 F	Refine search	pes		Interval 443-636 2159-236	E-val 2.800 9 6.020
List of domai	Accession pfam01365 pfam01365 cd12878	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2)	Search for sin	nilar domain architectur Descripti nology) domain is an eo nology) domain is an eo septor (RvR); This SPR	es 2 F	Refine search	rpes rpes ond of		Interval 443-636 2159-2369 1072-120	E-val 2.80e 9 6.02e 4 1.78e
List of domai Name RYDR_ITPR RYDR_ITPR SPRY2_RyR RR TM4-6	n hits Accession pfam01365 pfam01365 cd12878 pfam06459	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM 4	Search for sin R and IP3R Hor R and IP3R Hor of ryanodine re- 4-6; This region	Descripti Descripti nology) domain is an ex nology) domain is an ex exptor (RyR); This SPR covers TM regions 4-6	on tracellular don tracellular don Y domain (SPF of the ryanodin	Refine search main from two ty main from two ty RY2) is the sector e receptor 1	rpes rpes ond of		Interval 443-636 2159-2361 1072-120- 4383-467	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e
ist of domai Name RYDR_ITPR RYDR_ITPR SPRY2_RyR RR_TM4-6 SPRY1_RyR	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM - SPRY domain 1 (SPRY1)	Search for sin R and IP3R Hor R and IP3R Hor of ryanodine rei 4-6; This region of ryanodine rei	Descripti Descripti nology) domain is an ex pology) domain is an ex poptor (RyR); This SPR covers TM regions 4-6 septor (RyR); This SPR	on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain is the	Refine search main from two ty main from two ty RY2) is the sect e receptor 1 e first of three	Pes pes ond of		Interval 443-636 2159-236 1072-120 4383-467 642-793	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e
List of domai Name RYDR_ITPR RYDR_ITPR RYDR_ITPR RPY2_RyR RR_TM4-6 SPRY1_RyR SPRY3_RyR	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879	RiH domain; The RiH (Ry RiH domain; The RiH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM SPRY domain 1 (SPRY1) SPRY domain 3 (SPRY3)	Search for sin R and IP3R Hor R and IP3R Hor of ryanodine re- of ryanodine re- of ryanodine re-	Descripti Descripti nology) domain is an ex nology) domain is an ex septor (RyR); This SPR covers TM regions 4-6 .septor (RyR); This SPR septor (RyR); This SPR	es 2 F on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain is the Y domain (SPF	Refine search main from two ty nain from two ty RY2) is the secc e receptor 1 e first of three RY3) is the third	2 rpes pes ond of		Interval 443-636 2159-2363 1072-120 4383-467 642-793 1418-156	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e
List of domai Name RYDR_ITPR BRYDR_ITPR BSPRY2_RyR RR_TM4-6 SPRY3_RyR Ins145_P3_rec	Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam08709	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM SPRY domain 1 (SPRY1) SPRY domain 3 (SPRY3) Inositol 1,4-5risphosphat	Search for sin R and IP3R Hor of ryanodine ree 4-6; This region of ryanodine ree te/ryanodine rec	bilar domain architectur Descripti nology) domain is an er nology) domain is an er sceptor (RyR); This SPR covers TM regions 4-6 covers TM regions 4-6 covers TM regions 4-6 covers TM regions 4-6 covers TM regions 4-6 regions 4-7 regions 4-6 regions 4-7 regions 4-7	es ? F on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain (SPF esponds to the	hain from two ty nain from two ty RY2) is the seco e receptor 1 e first of three RY3) is the third ligand binding	Pes pes ond of		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-1566 8-203	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e 5.34e
List of domai Name PRYDR_ITPR RYDR_ITPR SPRY2_RyR RR_TM4-6 SPRY1_RyR SPRY3_RyR Ins145_P3_rec MIR	n hits Accession pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam08709 pfam08709 pfam02815	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM - SPRY domain 1 (SPRY1) SPRY domain 3 (SPRY2) Inositol 1,4,5-trisphosphal MIR domain; The MIR (pr	Search for sin R and IP3R Hor R and IP3R Hor of ryanodine re- of ryanodine re- of ryanodine re- of ryanodine re- of ryanodine rec- otein mannosylt	nilar domain architectur Descripti nology) domain is an ey apptor (RyR); This SPR covers TM regions 4-6 reptor (RyR); This SPR ceptor (RyR); This SPR petor; This domain corr ransferase, IPSR and R	es 2 F on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain (SPF Y domain is th Y domain (SPF Y domain is th Y domain is th y domain is th y domain is th y domain is to the	hefine search nain from two ty anain from two ty ty?2) is the secc e receptor 1 e first of three ty?3 is the third ligand binding a domain that m	pes pes ond of ond nay		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e 5.34e 1.70e
List of domai Name RYDR_ITPR RYDR_ITPR SPRY2_RyR SPRY1_RyR SPRY1_RyR SPRY3_RyR Ins145_P3_rec MIR MIR	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam08709 pfam02815 pfam02026	RIH domain; The RIH (Ry RIH domain; The RIH (Ry SPRY domain 2 (SPRY2) SPRY domain 1 (SPRY1) SPRY domain 3 (SPRY3) Inositol 1,4,5-trisphosphal MR domain; The MR (pr RyR domain; This domain	Search for sin R and IP3R Hor R and IP3R Hor of ryanodine re- de-6; This region of ryanodine re- tel/ryanodine re- tel/r	Initiar domain architectur Descripti nology) domain is an ex- petor (RyR); This SPR overs TM regions 4-6 ceptor (RyR); This SPR septor (RyR); This SPR septor; This domain corr ansferase, IP3R and R r Ryanodine receptor.	es 2 F on tracellular don tracellular don tracellular don y domain (SPF of the ryanodin Y domain is the Y domain (SPF esponds to the esponds to the apponds to the first of the transmission (SPF) appondent of the sponds to the the domain is in the domain is in	Refine search main from two ty nain from two ty RY2) is the secc e receptor 1 RY3) is the third ligand binding a domain that n found in four	Ppes ppes ond of of nay		Interval 443-636 2159-236 1072-120 4333-467 642-793 1418-156 8-203 211-389 850-940	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.000 5.34e 1.70e 1.97e
List of domai Name RYDR_ITPR RYDR_ITPR RYDR_ITPR RTM4-6 SPRY1_RyR SPRY3_RyR Ins145_P3_rec MIR RyR RyR	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12877 pfam08709 pfam08709 pfam02815 pfam02026	RiH domain: The RiH (Ry RiH domain: The RiH (Ry SPRY domain 2 (SPRY2) Ryanodine Receptor TM SPRY domain 1 (SPRY1) SPRY domain 1 (SPRY1) MR domain: The dMR (gr RyR domain: The domain RyR domain: This domain	Search for sir R and IP3R Hor R and IP3R Hor of ryanodine re- dr-6; This region of ryanodine re- te/ryanodine re- te/ryanodine re- te/ryanodine re- totein manosyti i is called RyR fr	nilar domain architectur Descripti nology) domain is an er ology) domain is an er zeptor (RyR); This SPR zeptor (RyR); This SPR zeptor (RyR); This SPR petor; This domain corr ansferase, IP3R and R or Ryanodine receptor. r Ryanodine receptor.	es 2 F on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain (SPF y domain is (SPF esponds to the y(R) domain is i The domain is i	Refine search main from two ty nain from two ty RY2) is the secc e receptor 1 RY3) is the third ligand binding a domain that m found in four found in four	Ppes ppes ond of of nay		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389 850-940 2735-282	E-val 2.80e 9 6.02e 4 1.78e 1.03e 6 4.03e 5.34e 1.70e 5.34e 1.70e 5.31e 1.97e 5 1.25e
List of domai Name HYDR_ITPR HYDR_ITPR HSPR72_RYR HSPR71_RYR HSPR71_RYR HSPR73_RYR HIN146 HSPR71_RYR HSPR74_RY	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam08709 pfam028709 pfam02026 pfam02026	RIH domain; The RIH (Ry RIH domain; The RIH (Ry Ryanodine Receptor TM SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain; The MIR (pr RyR domain; The domain RyR domain; This domain RyR domain; This domain	Search for sir R and IP3R Hor R and IP3R Hor R and IP3R Hor of ryanodine re of ryanodine re of ryanodine rec of ryanodine rec otein mannosylt is called RyR fr is called RyR fr	Initiar domain architectur Descripti nology) domain is an ei peptor (RyR); This SPR overs TM regions 4-6 covers TM regions 4-6 covers TM regions 4-6 peptor (RyR); This SPR optor, This domain corr ransferase, IPSR and R or Ryanodine receptor. Ir Ryanodine receptor.	es 2 F on tracellular don tracellular don Y domain (SPF of the ryanodin Y domain is the Y domain is the Y domain is the Y domain is in The domain is in The domain is in The domain is	Aefine search main from two ty hain from two ty Y(2) is the secc e receptor 1 e first of three Y(3) is the thirding a domain that in found in four found in four found in four	2 pes pes nd of		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389 850-940 2735-282 964-1054	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e 5.34e 1.97e 1.97e 5 1.25e 4 6.92e
List of domail Name RYDR_ITPR RYDR_ITPR SPRY2_RYR SPRY2_RYR RSTN_4-6 SPRY1_RYR SPRY3_RYR Ins145_P3_rec MIR RyR RyR RyR RyR	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12877 cd12877 pfam02815 pfam02026 pfam02026 pfam02026	RiH domain: The RiH (Ry SPRY domain 2 (SPRY) SPRY domain 2 (SPRY) SPRY domain 3 (SPRY3) Inositol 1,4.5-trisphosphat MR domain: The MR (pr RyR domain; This domain RyR domain; This domain RyR domain; This domain	Search for sir R and IP3R Hor of ryanodine re- of ryanodine re- of ryanodine re- of ryanodine re- of ryanodine re- tel/ryanodine re- tel/ryanodine re- tel/ryanodine re- nis called RyR fr is called RyR fr	Descripti nology i domain is an ex- nology j domain is an ex- nology j domain is an ex- poptor (RyR). This SPR covers TM regions 4-6 exptor (RyR). This SPR septor (RyR). This SPR pelor, This domain corr ansferase, IP3R and R r Ryanodine receptor. r Ryanodine receptor.	es 2 F on tracellular don tracellular don Y domain (SPF esponds to the esponds to the yR) domain is The domain is The domain is The domain is The domain is	Refine search main from two ty nain from two ty 742) is the secc e receptor 1 first of three Y33) is the third ligand binding a domain that n found in four found in four found in four	2 pes pes nd of		Interval 443-636 2159-236 1072-120 4383-467 642-793 211-389 850-940 2735-282 964-1054 2855-293	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e 5.34e 1.70e 5 1.25e 6.92e 9 8.71e
List of domai Name PRVDR_ITPR BRVDR_ITPR BSRV2_RyR RR_TM4-6 BSRV1_RyR BSRV1_RyR BSRV3_RyR BSRV3_RyR BVR BVR BVR BVR BVR BVR BVR BVR BVR BV	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam06459 cd12877 pfam02455 pfam02026 pfam02026 pfam02026 smart00449	RIH domain; The RIH (Ry RIH domain; The RIH (Ry Ryanodine Receptor TM SPRY domain 3 (SPRY1) SPRY domain 1 (SPRY1) SPRY domain; The MR (ry Ry domain; The domain RyR domain; The domain	Search for sir R and IP3R Hor R and IP3R Hor of ryanodine re- of ryanodine re- of ryanodine re- tefryanodine rec- tefryanodine rec- tefryanodine rec- tefryanodine rec- tefryanodine rec- tefryanodine rec- tefryanodine rec- ter mannosylt is called RyR fr is called RyR fr Yanodine Rece-	Descripti Descripti nology) domain is an ex- nology) domain is an ex- poptor (RyR); This SPR covers TM regions 4-6 reptor (RyR); This SPR peptor, This domain corr ansferase, IP3R and R r Ryanodine receptor. r Ryanodine receptor. r Ryanodine receptor. r Ryanodine receptor.	es 2 F on tracellular don tracellular don tracellular don tracellular don y domain (SPF of the ryanodin Y domain (SPF Y domain is The domain is The domain is The domain is The domain is n function. Dis	Refine search main from two by nain from two by RY2) is the secc e receptor 1 e first of three RY3) is the third ligand binding a domain that m found in four found in four found in four found in four found in four tant homologue	pes ppes nnd of of nnay		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389 850-940 2735-2822 964-1054 2855-293 1064-120 1064-120	E-val 2.800 9 6.020 4 1.786 1 3.366 1.036 6 4.000 5.346 1.706 1.976 5.346 1.706 1.976 5.988 4.6.922 9 8.716 6.988
List of domai Name RYDB_TIPR RYDB_TIPR SPRY2_RYR RTM4-6 SPRY1_RYR Ins145_P3_rec MIR RyR RyR I RyR I RyR I SPRY SPRY SPRY I SPRY I SPRX I SPRX I SPRX I SPRX I SPRX I SPRX I SPRX I SPRX I SPRX I SPRX	n hits Accession pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam02815 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026	RiH domain: The RiH (Ry SPRY domain 2 (SPRY2) SPRY domain 2 (SPRY2) SPRY domain 3 (SPRY3) Inositol 1,4,5-trisphosphal RiR domain: The MIR (pr RyR domain; This domain RyR domain; This domain Domain in SPIa and the F SPRY domain; SPRY Dor	Search for sir R and IP3R Hor R and IP3R Hor of ryanodine re- d-6; This region of ryanodine re- of ryanodine re- defryanodine re- defryanodine re- defryanodine re- te/ryanodine Re- nis called RyR fr is called RyR fr	Initer domain architectur Descripti nology) domain is an ex- nology) domain is nore peptor (RyR); This SPR covers TM regions 4-6 septor (RyR); This SPR captor (RyR); This SPR captor (RyR); This SPR peptor, This domain cor manferrase, IPSR and R ex- ptor, This domain cor respondence receptor; respondence rec	es 2 F on tracellular don tracellular don tracellular don tracellular don tracellular don y domain (SPF of the ryanodin Y domain is i The domain is i	Refine search main from two ty nain from two ty have the secce e receptor 1	Image: Second		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389 850-940 2735-282 964-1054 2855-293 1064-120 1066-120 1066-120	E-val 2.80¢ 9 6.02¢ 4 1.78¢ 1.03¢ 6 4.00¢ 5.34¢ 1.70¢ 1.97¢ 5 1.25¢ 6 6.92¢ 9 8.711¢ 6 6.98¢ 6 6.638¢
List of domai Name PRYDR_ITPR PRYDR_ITPR PSPR7_EXR PSPR7_EXR PSPR7_EXR PSPR7_EXR PSPR7_EXR PSPR7	n hits Accession pfam01365 cd12878 pfam01365 cd12877 cd12877 cd12877 pfam02876 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026	RiH domain; The RiH (Ry RiH domain; The RiH (Ry Ryanodine Receptor TM SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain; GPRY1) NR domain; The domain RyR domain; The domain RyR domain; The domain RyR domain; SPRY Do- SPRY domain; SPRY DO- S	Search for sir R and IP3R Hor R and IP3R Hor of ryanodine re- de; This region of ryanodine rec of ryanodine rec of ryanodine rec otein manosyti is called RyR ft is called RyR ft is called RyR ft is called RyR ft is called RyR ft rais called RyR ft rais called RyR ft associated; Thi associated; Thi	Illar domain architectur Descripti Inology) domain is an ex nology) domain is more peptor (RyR); This SPR optor (RyR); This SPR optor (RyR); This SPR optor (RyR); This SPR optor, This domain cor ansferase, IP3R and R Ryanodine receptor. R Kyanodine receptor. R Kyanodine receptor. R Kyanodine receptor.	es 2 F F on tracellular don tracellular don Y domain (SPY 4 domain is th Y domain is th Y domain is th Y domain is th Y domain is th P domain is The domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain is the domain	Refine search main from two by main from two by RY2) is the secc e receptor 1 Sinst of three RY3) is the third ligand binding a domain that n found in four found in four found in four found in four found in four tant homologue Domain of un line receptors (i Domain of in	Impes ppes ond of and of <t< td=""><td></td><td>Interval 443-635 2159-236 1072-120 4383-467 642-793 31418-156 8-203 211-389 850-940 2735-282 964-105 2855-293 1084-120 1084-100 1084-100 1084-100 1</td><td>E-val 2.80e 9.6.02e 4.1.78e 1.336e 1.03e 6.4.00e 5.34e 1.70e 5.1.25e 6.92e 9.8.71e 6.92e 9.8.71e 6.63e 2.58e 5.34e 5.3</td></t<>		Interval 443-635 2159-236 1072-120 4383-467 642-793 31418-156 8-203 211-389 850-940 2735-282 964-105 2855-293 1084-120 1084-100 1084-100 1084-100 1	E-val 2.80e 9.6.02e 4.1.78e 1.336e 1.03e 6.4.00e 5.34e 1.70e 5.1.25e 6.92e 9.8.71e 6.92e 9.8.71e 6.63e 2.58e 5.34e 5.3
List of domai Name H RVDR_ITPR H RVDR_ITPR H RVDR_ITPR H RCT_ITA4.6 H SPRY1_RVR H SPRY3_RVR H Ins145_P_roc H MIR H RVR H RVR H RVR H SPRY H SPRY H SPRY H RVR H SPRY H RVR H SPRY H RVR H SPRY H RVR H SPRY H RVR H RVR H SPRY H RVR H RVR H SPRY H RVR H RVR H SPRY H RVR H RVR H RVR H RVR H RVR H SPRY H RVR H RVR H RVR H RVR H RVR H RVR H	n hits Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 pfam02815 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam02026 pfam00622 pfam06520	RIH domain: The RIH (Ry SPRY domain 2 (SPRY2) SPRY domain 2 (SPRY2) SPRY domain 3 (SPRY3) Inositol 1,4,5-trisphosphat RIR domain; The domain RyR domain; This domain RyR domain; This domain Domain in SPIa and the F RyR domain; This domain Domain in SPIA and HPAR Homology SPRY domain; SPRY Dor RyR and IPAR Homology RyR domain; SPRY Dor	Search for sir R and IP3R Hor R and IP3R Hor of ryanodine re teryanodine re teryanodine ro teryanodine ro teryanodine ro teryanodine ro teryanodine ro teryanodine ro teryanodine ro teryanodine and nis called RyR fr is called RyR fr	Inlar domain architectur Descripti nology (domain is an ex- nology) (domain is an ex- nology) (domain is an ex- nology) (domain is an ex- patra (RyR); This SPR ouvers TM regions 4-5- aptor (RyR); This SPR optor, This domain cor- pator (RyR); This SPR optor, This domain cor- pator (RyR); This SPR optor, This domain cor- r Ryanodine receptor. F Ryanodine receptor. F Ryanodine receptor. Tor, Domain of unknow om SPla and the RYana om SPla and the RYana	on tracellular don tracellular don tracellular don tracellular don Y domain (SPF y domain is the Y domain is the Y domain is in The domain is The d	Refine search hain from two ty hain from two ty hain from two ty receptor 1 first of three erst of three erst of three erst of three erst of three three the third ligard binding a dmain that in dourd in four found in four found in four found in four tant homologue Domain of un ine receptors (Domain of un ponele. Thie	pes rpes nnd of hay known RyR) known		Interval 443-636 2159-236 1072-120 4383-467 642-793 1418-156 8-203 211-389 850-940 2735-282 964-1054 2855-293 1086-1200 1086-1200 1086-1200 1086-200 3879-399 660-755 7656-755	E-val 2.80e 9.6.02e 4.1.78e 1.03e 1.03e 5.34e 1.70e 1.97e 5.1.25e 4.6.92e 9.8.71e 6.6.98e 6.6.88e
List of domai Name H RVDR_ITPR H SPDR_ITPR H SPRVT_RVR H SPRVT_RVR H SPRVT_RVR H SPRVT_SVR H SPRVT_SVR H SPRY H SPRY	n hits Accession pfam01365 cd12878 pfam064278 cd12879 cd12877 cd12877 cd12877 cd12877 cd12877 cd12877 pfam02026 pfam02026 pfam02026 pfam00222 pfam04454 pfam00520 pfam00520	RiH domain; The RiH (Ry RiH domain; The RiH (Ry Ryanodine Receptor TM SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain; GPRY1) SPRY domain; The domain RyR domain; The domain RyR domain; The domain RyR domain; SPRY Do SPRY domain; SPRY Do	Search for sir R and IP3R Hor R and IP3R Hor R and IP3R Hor of ryanodine re- 4-6; This region of ryanodine re- tef/ryanodine re- of ryanodine re- tef/ryanodine Re- nis called RyR fi nis called RyR fi response to the second file of the second file of the second file of the second file	Illar domain architectur Descripti Indiogy) domain is an ex- nology) domain is an ex- polyr (RyR); This SPR operor	es 2 F on tracellular don tracellular don tracellular don tracellular don tracellular don tracellular don y domain is bir tracellular don tracellular	Refine search hain from two ty hain from two ty AV2) is the seco e receptor 1 of first of three YY3) is the third ligand binding a domain that n found in four found in four	Pess ppess ond of </td <td></td> <td>Interval 443-635 2159-236 1072-120 4383-467 642-793 211-389 850-940 2735-282 964-1054 2855-293 1084-120 1084-1084-100 10</td> <td>E-val 2.80c 9 6.02e 4 1.78e 1 3.36e 1.37e 5 1.25e 4 6.92e 9 8.71 5 6.92e 9 8.71 6 6.98e 6 6.63e 2 2.58e 7.17e 9 1.68e</td>		Interval 443-635 2159-236 1072-120 4383-467 642-793 211-389 850-940 2735-282 964-1054 2855-293 1084-120 1084-1084-100 10	E-val 2.80c 9 6.02e 4 1.78e 1 3.36e 1.37e 5 1.25e 4 6.92e 9 8.71 5 6.92e 9 8.71 6 6.98e 6 6.63e 2 2.58e 7.17e 9 1.68e
List of domai Name H RVDR, ITPR H RVDR, ITPR H SPRY_EYR B RF_TM4-6 H SPRY1_FyR H SPRY1_FyR H SPRY5 H SPRY H SPR	n hits Accession pfam01385 cd12878 cd12877 cd12877 pfam08709 pfam02875 pfam02026 pfam02026 pfam02026 pfam02026 pfam020206 pfam082020 pfam00622 pfam08622 pfam08622 pfam08622	RIH domain: The RIH (Ry SPRY domain 2 (SPRY2) SPRY domain 2 (SPRY2) SPRY domain 3 (SPRY3) Inositol 1,4.5-trisphosphat RIR domain: The domain RyR domain; This domain RyR domain; This domain Domain in SPIs and the F RyR domain; This domain Domain in SPIs and the F RyR domain; SPRY dom RyR and IP3R Homology RyR domain; SPRY dom RyR and IP3R Homology RYR domain; SPRY dom RyR and IP3R Homology RYR domain; SPRY dom	Search for sil R and IP3R Hora R and IP3R Hora d Tyanoline re- 46; This region d Tyanoline re- te/syanoline re- te/syanoline re- te/syanoline re- colien mannosyth is called RyR fr is called RyR	Initial domain architectur Descripti nology / domain is an ex- nology / domain is an ex- nology / domain is an ex- nology / domain is an ex- powers TM regions 4-6- septor (RyR): This SPR optor, This domain core petor, RyR, This SPR aptor (RyR): This SPR aptor (RyR	on tracellular don tracellular don tracellular don tracellular don tracellular don tracellular don ty domain (SPP domain (SPP domain (SPP domain S) the domain is the domain i	teline search hain from two ty hain from two ty hain from two ty Y2(2) is the secce a receptor 1 first of three Y2(3) is the third ligand binding a domain that n dound in four found in four	Pees rpes rpes nod of		Interval 443-636 2159-236 1072-120 4383-467 642-757 642-757 8203 211-389 850-946 965-1052 265-210 1084-120 1084-120 3679-396 660-7854 1431-1565	E-val 2.80e 9 6.02e 4 1.78e 1 3.36e 1.03e 6 4.00e 5.34e 1.70e 5 1.25e 6 6.92e 9 8.71e 6 6.92e 9 8.71e 6 6.92e 9 8.71e 8 6.92e 9 8.81e 8 6.92e 9 8.81e 8 6.92e 9 8.81e 8 6.92e 9 8.71e 8 6.92e 9 8.81e 8 6.92e 9 8.71e 8 6.92e 9 8.81e 8 8.81e 8 8.92e 9 8.81e 8 8.81e 8 8.92e 9 8.81e 8 8.81e 8 8.92e 9 8.81e 8 8.92e 9 8.85e 8 8.81e 8 8.92e 9 8.85e 8 8.85e 8 8.92e 9 8.85e 8 8.85e 8 8.92e 9 8.85e 8 8.92e 8 9.92e 8 9.
List of domai Name + RYDR ITPR + RYDR ITPR + SPRYZ_RYR + SPRYZ_RYR + Inst45_P3_rec + MIR + SPRY + S	n hits Accession pfam01385 of1278 pfam01385 of12878 pfam02459 pfam0270 pfam0270 pfam0270 pfam02702 pfam022026 pfam02026 pfam02026 pfam02022 pfam08454 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520	RiH domain; The RiH (Ry RiH domain; The RiH (Ry Ryanodine Receptor TM SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain; ISPRY10 Inostal 1,4,5-14;SPRY30 Inostal 1,4,5-14;SPRY30 RyR domain; Tha domair RyR domain; Tha domair RyR domain; SPRY Do SPRY domain; SPRY Do	Search for sill R and IP3R Hord R and IP3R Hord of vyanodine re- dof vyanodine re- si called RyR fi is called RyR fi vanobia re- main is named fi vanobia re- ventine Re- ventine Re- ventine Re- ventine Re-	Illar domain architectur Descripti Inology) domain is an ex nology) domain is more peptor (RyR); This SPR operor (on traceolular don traceolular don tr	terine search hain from two ty hain from two ty YZ) is the sect e receiptor 1 first of three YZ) is the third ligand binding a domain that i e receiptor 1 found in four found in four	Impes ppes ond of i of may ks are kknown RyR) RyR) ss are known ss are		Interval 443-636 2159-236 1072-120 4383-467 642-793 211-389 850-940 2735-282 964-1054 2855-293 1086-1200 1080-1080-	E-val 2.80e 9.6022 4.1.78e 1.03e 5.34e 1.70e 1.97e 5.4.25e 8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.71e 9.8.72e 9.8.71e 9.8.71e 9.8.72e 9.8.71e 9.8.71e 9.8.72e 9.8.71e 9.8.72e 9.8.71e 9.8.72e 9.8.72e 9.8.71e 9.8.72e
List of domai Name + RYDR_ITPR + RYDR_ITPR + SPRV1_FPR + SPRV2_FYR + SPRV1_FYR + SPRV3_FYR + SPRV3_FYR + SPRV4 + SPR	Inits Accession pfam01385 pfam01385 pfam01385 pfam04459 pfam022815 pfam022026 pfam02026 pfam08429 pfam02026 pfam08000 pfam080000 pfam080000 pfam080000 pfam080000 pfam080000 pfam080000 pfam080000 pfam0800000 pfam08000000 pfam0800000000000000000000000000000000000	RIH domain: The RIH (R; SPRY domain 2 (SPRY2) SPRY domain 2 (SPRY2) SPRY domain 3 (SPRY3) Inositol 1.4.5-trisphosphal SPRY domain 3 (SPRY3) Inositol 1.4.5-trisphosphal MR domain: The domain RyR domain; This domain RyR domain; This domain Domain in SPIa and the R SPRY domain; SPRY Dor RyR and IP3R Homology SPRY domain; SPRY Dor Domain in SPIa and the R Domain in SPIa and the R Domain in SPIa and the R	Search for sill R and IP3R Hor R and IP3R Hor R and IP3R Hor of yanodine re- 46: This region of yanodine re- cototin mannsytt is called RyR R is is called RyR R is called RyR R is called RyR R yanodine Rece region and the region of the region sesociated: This rame of the region of the region of the region the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the region of the regi	nilar domain architectur Descripti nology I domain is an ex- nology I domain is an ex- nology I domain is an ex- nology I domain is an ex- peter (FN) is the step- aptor (FN). This SPR aptor (FN). This SPR aptor FX and FR and R F Nanodine receptor. F Nanodine receptor. F Nanodine receptor. F Nanodine receptor. F Nanodine receptor. Data SPI and the RYan eukaryoti domain is di Nano SPI and the RYan on SPI and the RYan or SPI and the RYan or SPI and the FYan for Domain of unknow	es 2 F tracellular don tracellular don tracellular don tracellular don tracellular don tracellular don ty domain (SPF 4 domain (SPF 4 domain (SPF 4 domain Si The domain is the domain is th	teline search hain from two ty hain from two ty hain from two ty Y(2) is the sect e receptor 1 first of three Y(2) is the third ligand binding a domain that n found in four found in four found in four found in four found in four that homologue Domain of un annels. This Domain of un tan thomologue tant homologue tant homologue	Image: Second		1443-353 2159-236 1077-20- 4383-407 4383-407 4382-407 2132-80- 367-36 367-36 367-36 367-36 367-36 367-36 367-36 367-36 367-36 367-36 367-36 36 377-36 36 377-36 36 377-36 36 377-36 36 377-36 36 377-36 36 377-377-377-377-377-377-377-377-377-377	E-val 2.80e 9 6.02e 4 1.78e 1.03a 6 4.00e 5.34e 1.70e 1.97e 5 1.25e 6 6.92e 9 8.71e 6 6.92e 9 8.71e 6 6.92e 9 8.71e 6 6.92e 9 8.71e 8 6.92e 9 8.62e 8 6.92e 9 8.62e 8 6.92e 9 8.62e 8 6.92e 9 8.62e 8 6.92e 9 7.77e 8 6.92e 9 7.72e 8 6.92e 9 7.72e 8 6.92e 9 7.72e 8 6.92e 9 7.72e 8 6.92e 9 7.72e 8 6.92e 8 6.92e 9 7.72e 8 6.92e 8 7.72e 8
List of domail Name RYDR, ITPR BYDR, ITPR BYDR, ITPR BYDR, ITPR BYRT, IFR BYRT, IFR BYRT, IFR BYRT, IFR BYRT, IFR BYRT BYRT BYRT BYRT BYRT BYRT BYRT BYR	n hits Accession pfam01385 of 201287 pfam01385 of 201287 pfam06459 of 2012877 of 2012879 pfam02028 pfam02028 pfam02028 pfam02028 pfam02022 pfam08454 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam00520 pfam0520	RiH domain; The RiH (Ry RiH domain; The RiH (Ry Ryanodine Receptor TM SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain (SPRY1) SPRY domain; The domain RyR domain; The domain RyR domain; The domain RyR domain; The domain RyR domain; SPRY Dom SPRY dom S	Search for sil R and IP3R Hot R and IP3R Hot R and IP3R Hot Hot R and IP3R Hot of yanodine re- do yanodine re- do yanodine re- do yanodine re- do yanodine re- to di yanodine re- to di yanodine re- to called RYR fi is called RYR fi yanodine Rece Yanodine Rece Yanodine Rece	Illar domain architectur Descripti Indigo) domain is an ex- nology) domain is an ex- nology) domain is an ex- petror (RyR). This SPR operor (RyR). This SPR operating (RyR). This SPR	es 2 F F on tracellular don tracellular don tracellul	terine search hain from two ty hain from two ty Y2) is the sect e receptor 1 first of three Y2) is the third ligand binding a domain that i v2) is the third ligand binding a domain that i domain for found in four found in four four four four four terms four four four four terms four	τ τ τ τ		Interval 443-635 2159-238 11072-120 4323-467 482-733 211-389 85-9-94 273-5282 2855-233 1084-120 1086-120 378-399 866-795 4765-494 1433-156 866-195 4765-494 1433-156 866-195 4765-494 2475-25	E-val 2.806 9 6.022 4 1.786 1.036 6 4.006 5.346 1.706 1.976 5 1.256 6 6.986 6 6.638 6 6.6382 2 2.586 7.176 9 1.688 8 4.816 8 3.226 6.044 3 5.822

CDD provides a guide to domain level architecture, including sequence alignments & representative structures.

			<u>}</u>	
E SEARCH G		rch Structure Home 3D Macromolecular Structures Conserved Domains	Pubchem B	lioSyster
		New York (1971)	Other devid Devid	
		Jn [sp [P2181/] view sater 1 OS-Home spring OV-0606 CN-DVD1 DE-1 SV-3 view	Standard Resul	its 🗸 i
IKI_HUMAN	Ryanounie rec			
Protein Clas	ssification	nain-containing protein (domain architecture ID 11606398)		
otein containin	g domains RYDR	ITPR, SPRY1_RyR, SPRY2_RyR, and RR_TM4-6		
Graphical s	ummary 🔲	Zoom to residue level show extra options »		(
	1	1000 2000 3000 4000	5438	
uery seq. necific bits				
peerire nrea	· • • • • • • • • • • • • • • • • • • •	SPRY SPR SPRY	Ion_t	
		SPR SPR SPRY		
on-specific	In			
uperfamilies	I I RY	R_ SPRY Ry Ry SPR SPRY RYDR_I Ry Ry RH E RR_TH4-6	Ion_t	
	MIR s			
		Search for similar domain architectures Refine search		
ist of dom	ain hits			
Name	Accession	Description	Interval	E-value
RYDR_ITPR	pfam01365	RIH domain; The RIH (RyR and IP3R Homology) domain is an extracellular domain from two types RIH domain: The RIH (RyR and IP3R Homology) domain is an extracellular domain from two types	443-636 2159-2369	2.80e-8 6.02e-8
SPRY2_RyR	cd12878	SPRY domain 2 (SPRY2) of ryanodine receptor (RyR); This SPRY domain (SPRY2) is the second of	1072-1204	1.78e-8
RR_TM4-6	pfam06459	Ryanodine Receptor TM 4-6; This region covers TM regions 4-6 of the ryanodine receptor 1	4383-4671	3.36e-8
SPRY1_RyR	cd12877	SPRY domain 1 (SPRY1) of ryanodine receptor (RyR); This SPRY domain is the first of three	642-793	1.03e-7
SPRY3_RyR	cd12879	SPRY domain 3 (SPRY3) of ryanodine receptor (RyR); This SPRY domain (SPRY3) is the third of	1418-1566	4.00e-7
Ins145_P3_re	c pfam08709	Inositol 1,4,5-trisphosphate/ryahodine receptor; This domain corresponds to the ligand binding	8-203	5.34e-7
	pfam02815	Mik domain; The Mik (protein mannosytiransierase, iPok and kyk) domain is a domain that may	211-389	1.708-7
RvR	pfam02026	RyR domain: This domain is called RyR for Ryanodine receptor. The domain is found in four	2735-2825	1.25e-4
RyR	pfam02026	RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four	964-1054	6.92e-3
RyR	pfam02026	RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four	2855-2939	8.71e-3
SPRY	smart00449	Domain in SPIa and the RYanodine Receptor; Domain of unknown function. Distant homologues are	1084-1206	6.98e-3
SPRY	pfam00622	SPRY domain; SPRY Domain is named from SPIa and the RYanodine Receptor. Domain of unknown	1086-1206	6.63e-3
RIH_assoc	pfam08454	RyR and IP3R Homology associated; This eukaryotic domain is found in ryanodine receptors (RyR)	3879-3992	2.58e-3
SPRY	pfam00622	SPRY domain; SPRY Domain is named from SPIa and the RYanodine Receptor. Domain of unknown	660-795	7.17e-2
lon_trans	pfam00520	ion transport protein; This family contains sodium, potassium and calcium ion channels. This	4765-4949	1.68e-2
determines ion membrane.	n selectivity. In sor	te sub-families (e.g. Na channels) the domain is repeated four times, whereas in others (e.g. K channels) the protein for Pssm-D: 334124 [Multi-domain] Cd Length: 237 Bit Score: 107.36 E-value: 1.68e-25	ns as a tetramer i	in the
		10 20 30 40 50 60 70 80		
arin	1917 474			
Cdd:	ofam00520 40	UFTGIFTLEMILKIIAAGFKITVLRSPwnildfvvvlpSLISLVLSSVGSLSGIRVLRILRVLRLLRLIRLEGLRTLVN 119		
		90 100 110 120 130 140 150 160		
		*		
		SVTHNGKOLVMTVGLLAVVVVLVTVVAFNFFRKFYNKSEDEDEPDMKCDDMMTCYLFHMYvgVRAGGGIGDEIEDPAGDE 4909		
sp Pi	21817 4830	,		
sp P; Cdd:j	21817 4830 pfam00520 120	SLIRSLKSLGNLLLLLLFLFIFAIIGYQLFGGKFYTWENPDNGRTNFDNFPNAFLWLFQTMTTEGWGDILYDTIDGK 197		
sp P: Cdd:p	21817 4830 pfam00520 120	SLIRSLKSLGNLLLLLLFFFFAIIGYQFFGGKFYTWENPONGRTNFDNFPNAFL#LFQTMTTECMGDILYDFIDGK 197		
sp P: Cdd:	21817 4830 ofam00520 120	SLIRSLKSLGNLLLLLLFLFIFAIIGYQLFGGKFYTWENPDNGRTNFDNFPNAFLMLFQTWTTEGMCDILYDTIDGK 197 170 180 190 200 *		
sp P; Cdd: sp P; Cdd:	21817 4830 pfam00520 120 21817 4910 pfam00520 198	SLIRSLKSLGNLLLLLLFLFFFAIIGVQLFGGKYYYMENPONGRYNFONPONAFLALFQTMTTEGAGDILVOTIDGK 197 170 180 200 170 170 197 200 VELKYWVDITFFFYVILLAIIGGLIIAAFGELAROOE 4949 GFWAIIYEFFIIGGELALAI-GUILONDERLEREE 237		
sp P; Cdd:r sp P; Cdd:r	21817 4830 ofam00520 120 21817 4910 ofam00520 198 ofam00520 198	SLIRSLKSLGNLLLLLLFLPIFAIIGVQLFGGKFYTWENPDNGRTNFDNFPNAFLALFQTHTTEGWGDILVOTIDGK 197 170 180 190 200 YELKRWYDDIFFFVIVILLALFIGULIDNFGELRNQQE 9499 GSFWAIITFVSFILGGFLLLMLFIGVIIDNFGELRNQDE 4949 GSFWAIITFVSFILGGFLLMLFIGVIIDNFGELRNQE 4949	1404 4500	4.94= 0
sp P; Cdd : [Cdd : [Cdd : [Cdd :]	21817 4830 ofam00520 120 21817 4910 ofam00520 196 pfam00622 smart00440	SLIRSLKSLGNLLLLLLFLPIFATIGYQLPGGKPYTWENPDNGRTNFDNFPNAFLKLFQTMTTEGMODILYDTIDGK 197 170 180 190 200 *	1431-1568	4.81e-2
sp P: Cdd : j Cdd : j Cdd : j SPRY SPRY	21817 4830 pfam00520 120 21817 4910 pfam00520 198 pfam00622 smart00449 smart00449	SLIBSLKSLGNLLLLLLFLFFIFATIGVQLFGGKFYTWENPDNGRTNFDNFPNAFLALFQTWTTEGWGDILVOTIDGK 197 170 180 190 200 YELKTWVPDTFFYUVILLALFIQULTIDNFGELADQG 6949 GSFWATIFVSFTILGGFLLGALFIQULTIDNFGELADQG 6949 SPRY domain SPR and the RYanodine Receptor: Domain of unknown Domain in SPR and the RYanodine Receptor: Domain of unknown function. Distant homologues are Domain in SPR and the RYanodine Receptor: Domain of unknown function. Distant homologues are Domain in SPR and the RYanodine Receptor: Domain of unknown function. Distant homologues are Domain in SPR and the RYanodine Receptor: Domain of unknown function. Distant homologues are	1431-1568 1430-1568 660-794	4.81e-2 3.22e-2 6.04e-1
sp P; Cdd : j Cdd : j Cdd : j - SPRY - SPR	21817 4830 pfam00520 120 21817 4910 pfam00520 198 pfam00622 smart00449 smart00449 pfam13833	SLIRSLKSLGNLLLLLLFLPIPATIGYQLPGGKPYTWENPDNGRTNPDNFPNAFLKLPQTMTTEGNDDILYOTIDGK 197 170 180 190 200 *	1431-1568 1430-1568 660-794 4083-4133	4.81e-2 3.22e-2 6.04e-1 5.82e-0
sp P: Cdd : j Cdd : j Cdd : j - SPRY - SPRY - SPRY - SPRY - SPRY - MIR	21817 4830 fam00520 120 21817 4910 pfam00520 198 pfam00622 smart00449 smart00449 pfam13833 smart00472	SLIBSLKSLGNLLLLLLFLFFITATIGVQLFQGKFYYWENPDNGRTNFDNFPNAFLALFQTWTTEGWGDILYDTIDGK 197 170 180 190 200 YELFKWPDTFFYTVILLATLGULINAFGELRQUG 4949 GSFNATIFVSFILLGULALLTUVILINAFGELRQUG 4949 GSFNATIFVSFILLGULALLTUVILINAFGELRQUG 4949 GSFNATIFVSFILLGULALLTUVILINAFGELRQUG 4949 Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPB and the RYanodine Receptor; Domain of unknown function. Distant homologues are	1431-1568 1430-1568 660-794 4083-4133 210-263	4.81e-2 3.22e-2 6.04e-1 5.82e-0 8.98e-0

CDD provides a guide to domain level architecture, including sequence alignments & representative structures.

NCBI		Conserved Domains	\$27210 \$325
ME SEARCH GUID	E NewSea	rch ST STRATE VERTURE TRATE Conserved Domain	s Pubchem BioSysten
	domains	- on [sp]P21817] entre 1 OS=Homo seniors OX=9606 GN=8VB1 PE=1 SV=3	View Standard Results 🗘 🕻
Protein Class	ification		•
PRY1_RyR and R	R_TM4-6 do	nain-containing protein (domain architecture ID 11696388)	
Craphical cur	iomains RYDR	11PK, SPRT1_KYK, SPRT2_KYK, and KK_IM4-6	
arapincar sur		1980 2000 1980 4000 4000 4000 4000 4000 4000 4000 4	5400
Query seq. Specific hits	MIR		FR_TH Ion_5
Ion-specific nits Superfamilies	In I I RY	SR SPRY N N SPR SPRY PIG_1 N N SP	RR_1774-6 Ion_t
	140 9	Search for similar domain architectures 2 Refine search	
ist of domai	n hits		
RYDR_ITPR RYDR_ITPR SPRY2_RyR RR_TM4-6 SPRY1_RyR SPRY3_RyR Ins145_P3_rec	Accession pfam01365 pfam01365 cd12878 pfam06459 cd12877 cd12879 pfam08709	Life domain; The RIH (RyR and IP3R Homology) domain is an extracellular domain from two types RIH domain; The RIH (RyR and IP3R Homology) domain is an extracellular domain from two types SPRY domain 2 (SPRY2) of ynanodine receptor (RyR); This SPRY domain (SPRY2) is the second of Ryanodine Receptor TM 46; This region covers TM regions 46 of the ryanodine receptor 1 SPRY domain 3 (SPRY3) of ryanodine receptor (RyR); This SPRY domain is the first of three SPRY domain 3 (SPRY3) of ryanodine receptor (RyR); This SPRY domain is the first of insolid 14.5-triberobschafter/ryanodine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/ryanodine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of insolid 14.5-triberobschafter/andine receptor (RyR); This SPRY domain (StPRY3) is the third of	Interval E-value 443-636 2.80e-8 4243-636 6.02e-8 1072-7804 1.78e-8 438-4671 3.36e-8 642-799 1.05-7 1418-1566 4.36e-7 8-203 5.34e-7
MIR RyR RyR RyR RyR RyR	pfam02815 pfam02026 pfam02026 pfam02026 pfam02026	MR domain; The MIR (protein mannesyltransferase, IP3R and RyR) domain is a domain that may RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four RyR domain; This domain is called RyR for Ryanodine receptor. The domain is found in four	211-389 1.70e-7 850-940 1.97e-4 2735-2825 1.25e-4 964-1054 6.92e-3 2855-2939 8.71e-3
SPRY SPRY RIH_assoc SPRY	smart00449 pfam00622 pfam08454 pfam00622	Domain in SPIa and the RYanodine Receptor, Domain of unknown function, Distant homologues are SPRY domain; SPRY Domain is named from SPIa and the RYanodine Receptor. Domain of unknown RyR and IP3R Homology associated; This eukaryotic domain is found in ryanodine receptors (RyR) SPRY domain; SPRY Domain is named from SPIa and the RYanodine Receptor. Domain of unknown In transport profile: This family contains softium notassium and calcium ign nahonder. This	1084-1206 6.98e-33 1086-1206 6.63e-3 3879-3992 2.58e-33 660-795 7.17e-21 4765-4949 168e-2
lon transport prot determines ion se membrane.	tein; This family electivity. In sor	contains sodium, potassium and calcium ion channels. This family is 6 transmembrane helices in which the last ne sub-families (e.g. Na channels) the domain is repeated four times, whereas in others (e.g. K channels) the pr	t two helices flank a loop which otein forms as a tetramer in the
		Pssm-ID: 334124 [Multi-domain] Cd Length: 237 Bit Score: 107.36 E-value: 1.68e-25	
sp P218 Cdd:pfa	817 4765 am00520 40	10 20 30 40 50 60 70 80 	4829 119
sp P218 Cdd:pfs	817 4830 1m00520 120	90 100 110 120 130 140 150 160 SVTHNGKGLWTVGLLAVVVITVVAFNFFKKYTWKSBEDEPDPKCCOMMTCYLFHKYvgVRAGGGGEDETEDFAGDE SLIKSLKSLGNLLLLLLFLFFAIIGYQLFGGKFYTWENPDKGRTHFDNFPNAFIMLFQTHTTEGMGDILYDTIDGK	4909 197
sp P218 Cdd:pfs	817 4910 1m00520 198	170 180 190 200 	
SPRY SPRY SPRY EF-hand_8	pfam00622 smart00449 smart00449 pfam13833	SPRY domain; SPRY Domain is named from SPIa and the RYanodine Receptor. Domain of unknown Domain in SPIa and the RYanodine Receptor; Domain of unknown function. Distant homologues are Domain in SPIa and the RYanodine Receptor; Domain of unknown function. Distant homologues are EF-hand domain pair;	1431-1568 4.81e-2 1430-1568 3.22e-2 660-794 6.04e-1 4083-4133 5.82e-0 040.22 0.02 0
Z INNES	ananu0472	comain in yanoune and nositor inspirospirate receptors and protein C-maintosyt/ansierases,	210-203 8.988-0

Once an initial trace is obtained for these regions, use DALI or PDBeFold to identify structural homologs that could not be identified by sequence alone.

- Start by identifying boundaries of conserved domains (NCBI CDD: <u>https://www.ncbi.nlm.nih.gov/Structure/cdd/</u>; DELTA-BLAST also performs CD-search by default)
- Then identify suitable structural templates for building known domains: FUGUE, SPARKS-X, PHYRE2, etc.
- Secondary structure, TM & disorder prediction (XtalPRED for overall summary; specific tools such as SPOT-DISORDER, SPIDER3 for best accuracy).
- Contact prediction from evolutionary couplings: EVFOLD & GREMLIN.
- Conservation analysis: Use favorite MSA algorithm (MUSCLE & CLUSTAL-OMEGA work well; TM-COFFEE, PRALINE-TM useful for membrane proteins) to create a sequence alignment of your protein with a few orthologs; gaps & insertions most commonly occur in loops/disordered regions. Useful as a guide during building.

XtalPRED is a great tool for summarizing predicted sequence properties.



Highlights predicted secondary structure, disorder, low complexity regions on sequence in an easily digestible format. Useful to print and consult while building. Also provides list of structural homologs. (http://ffas.burnham.org/XtalPred-cgi/xtal.pl)

(Also consider using some of the newer single purpose neural-network based classifiers; e.g. SPIDER-3 & SPOT-DISORDER-SINGLE from Yaoqi Zhou lab: <u>http://sparks-lab.org/index.php/Main/Services</u>)

Secondary structure prediction is a very useful guide when building.





Where is this motif in the sequence?

Secondary structure prediction is a very useful guide when building.



Secondary structure prediction is ~80% accurate. So if your model consistently disagrees with predicted secondary structure, look at it very closely!

What can we learn from the map alone?



What can we learn from the map alone?



Left handed! Obvious here – can be less clear at lower res, so be careful.

OK, that's better! What can we learn from the map alone?



Which direction does the helix point?



Which direction does the helix point?















Notice that the **absence** of large sidechain densities at small residue positions is just as valuable in validating the fit as the fit of large sidechains to the density.



Also, note that the information content of local regions varies. Consider "VTVVAASSTVV" vs "FGAAYWVTRA" – which is more likely to be uniquely identifiable from the map?



How to deal with uncertainty in sequence assignment and sidechain placement

- You will likely encounter situations where you cannot be certain of the local sequence register what to do?
- No clear consensus, but I suggest assigning residue code as "UNK" and numbering to "best guess" value. A more granular way to quantify/convey uncertainty would be helpful!
- Sidechain placement two main camps trim sidechains to density vs place them all (+/zero occ.). The former may sound more conservative, but it can hide errors during validation (during analysis of clashes). Either is acceptable, just be consistent, and preferably outline the approach taken when writing up the structure.

Prior knowledge can come in many forms – use any and all available info to guide model building.



Here, serendipitous identification of a conformational class of RyR1 lacking density for one subunit aided identification of protomer boundaries. In other cases, cross-linking data or NS data on subcomplexes or Fab-complexes may be helpful.

In a similar manner, we can use locally aligned difference maps between holo and apo structures to locate ligands.



The three ligands are clustered around the C-terminal domain.



The three ligands are clustered around the C-terminal domain.



Very good difference density even at moderate (3.8Å) resolution. Highlights importance of phases!



Secondary structure and the Ramachandran plot



- Describes geometrically favored backbone torsions (omega generally 180, except for prolines)
- At high resolution, outliers may be justified by density
- At low resolution, we can't see carbonyls, so much harder to justify Ramachandran outliers.
- This is the general-case Rama plot distribution is different for "special" residues (pro, gly, pre-pro)



(http://www.biochem.ucl.ac.uk/~martin/c40/peptide.html)

Helices – alpha and 3_{10}





Alpha

- ~90%
- 3.6 residues per turn
- Fat

310

- ~10%. More common in TM? (e.g. S4 of VSD)
- 3 residues per turn. Triangular cross section.
- Skinny
- Can be tricky to identify at low resolution, can lead to register errors.

Beta sheets

- Can be parallel or antiparallel in orientation (antiparallel more common and stable)]
- Twist of beta sheet varies, leading to more diversity than for alpha-helical structures.
- Harder to build at low resolution whereas a helix can be placed at ~7 Å, adjacent strands can only be clearly separated at ~4.5 Å.



(https://en.wikipedia.org/wiki/Beta_sheet)
EM-specific considerations

- No unambiguous sequence markers at low resolution (no equivalent of SeMet).
- No feedback from phase improvement, but also no model bias WYSIWIG.
- Often substantial variation in local resolution different strategies and levels of detail required for different regions. Map sharpening essential.
- "Medium" resolution (4-6Å) much more common than for crystallography.
- Often have more than one map, with different composition or conformation (combine focused refinements in Chimera by taking max value at each voxel after alignment, e.g.: vop maximum #1,2 ongrid #1)



Building an initial model - where to start?

- If you have a crystal structure, of a fragment or a homology model of a domain, place it, and extend into density.
- If you have sufficient resolution, try autobuilding with phenix
- Otherwise, identify structurally distinctive motifs in the sequence for example, a strongly predicted helix with three aromatic residues near the N-term end – and identify candidate locations in the density map. Extend and see if hypothesis still holds.



Start with map and model.



Move model to approximate position (if known, to save computation)



Run fitmap with 'search' (here 100 orientations) and 'radius' (here 5 Å)





Chimera will return a list of candidate orientations, ranked by agreement with the map. Hopefully there will be a clear separation between the correct and incorrect solutions.



Chimera will return a list of candidate orientations, ranked by agreement with the map. Hopefully there will be a clear separation between the correct and incorrect solutions.

Using UCSF Chimera for voxel size calibration (of your map and others)

- Voxel size generally requires calibration against a crystal structure.
- Once calibrated, generally stable between samples/datasets at same magnification.
- Can calibrate by fitting in Chimera at range of nominal voxel sizes and measuring correlation.
- Incorrect voxel sizes are common in deposited maps - be aware of this when comparing structures. E.g. here there is a 3% difference – affects structural alignment, reported resolution (3.8 vs 3.9Å).





- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



(Try the latest nightly with new features for EM, improved RSR: http://www.ccpem.ac.uk/download.php)



- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).

) 🛈 🖴 GitHul	b, Inc. (US) https://github.com/olibclarke/coot-tr	immings	C Search	☆ 自 ♥ ♣ 斋 ∢ 極→	+ ਡੁ≣	
Most Visited *	🛞 Read Later 😻 Getting Started 🛛 🖗 NEBc	utter - ORF S 🛞 Lab Book 🛞 🧱 ci	ryoSPARC 🎄 cryoSPARC (home)			
	Gilbclarke / coot-trimmings		⊙ Unwate	h 1		
	<>Code ① Issues 0 î\î Pul	I requests 0 III Projects 0	🗉 Wiki 🦾 Pulse 🔟 Graph	.s 🔅 Settings		
	Python customizations for the macromolecular model building software Coot. Add topics					
	B commits		♡ 0 releases	😃 1 contributor		
	Branch: master - New pull request		Create new file Uploa	ad files Find file Clone or download -	i	
	libclarke Fixed a bug in find_seque	nce		Latest commit dc0e68b 12 days ago		
	README.md	Update README.md		2 months ago		
	illi_custom.py	Fixed a bug in find_sequence.		12 days ago		
	I README.md					
	Coot-trimmings Python customizations for the macromolecular model building software Coot. Copy to ~/.coot-preferences (hidden dir, copy on command line) and restart coot. You should see a new menu ("Custom") and a bunch of new key bindings, as well as a couple of new toolbar buttons (0.0 "occurrence context")					



- Simple, intuitive interface for ٠ building and manipulating atomic models in density maps.
- Low computational ٠ requirements
- Extensive API easy to script or ٠ modify (using simple Python code)
- On-the-fly sharpening and low ٠ pass filtering (for MTZ).

••• 0	olibclarke/coot-trimmings: PV × +					
🗲 🛈 🔒 GitHub	p, Inc. (US) https://github.com/olibclarke/coot-trimm	nings C	C Search	☆自 ♥ ♣ 兪 ∢ ∞- ↔	'ਡ ≡	
Most Visited -	🛞 Read Later 🛛 🧕 Getting Started 🛛 🕵 NEBcutte	er - ORF S 🛞 Lab Book 🛞 🎇 cryo	oSPARC 🛛 🎇 cryoSPARC (home)			
	\mu olibclarke / coot-trimmings		O Unwatc	h ▼ 1 ★ Star 0 % Fork 0		
	Code Issues 0 DPull re	equests 0 III Projects 0 III	Wiki 🦇 Pulse 🛄 Graphs	s 🔅 Settings		
	Python customizations for the macr Add topics	omolecular model building softw	are Coot.	Edit		
	B commits		♡ 0 releases	🤽 1 contributor		
	Branch: master - New pull request		Create new file Uploa	d files Find file Clone or download -		
	olibclarke Fixed a bug in find_sequence	e		Latest commit dc0e68b 12 days ago		
	README.md	Update README.md		2 months ago		
	illi oli_custom.py	Fixed a bug in find_sequence.		12 days ago		
	I README.md					
	coot-trimmings					
	Python customizations for the					
	Copy to ~/.coot-preferences (Copy to ~/.coot-preferences (hidden dir, copy on command line) and restart coot.				
	You should see a new menu (' (e.g. "sequence context").	"Custom") and a bunch of new ke	ey bindings, as well as a cou	ple of new toolbar buttons		

Any Python (or Scheme) file you put in ~/.coot-preferences will be executed when starting Coot. Can use this for extra key bindings, scripts, custom functions.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).

def mutate_by_entered_code(): def mutate_single_letter(X): entry=str(X).upper() mol_id=active_residue()[0] ch_id=active_residue()[1] resno=active_residue()[2] ins_code=active_residue()[3] resname=residue_name(mol_id,ch_id,resno,ins_code) map_id=imol_refinement_map() aa_dic=('A':'ALA','R':'ARG','N':'ASN','D':'ASP','C':'CYS','E':'GLU','Q':'GLN','G':'GLY','H':'HIS','I':'ILE','L':'LEU','K':'LY nt_list=['A','C','T','G','U'] if (resname in aa_dic.values()) and (aa_dic.get(entry,0)!=0): mutate(mol_id,ch_id,resno,ins_code,aa_dic.get(entry,0)) elif (resname in nt_list) and (entry in nt_list): mutate_base(mol_id,ch_id,resno,ins_code,entry) else: info_dialog("Invalid target residue! Must be protein or nucleic acid, and entered code must be single letter.") generic_single_entry("New residue? (single letter code)","A","Mutate by single-letter code",mutate_single_letter)

#mutate active residue to entered residue code (upper or lower case single-letter)
add_key_binding("Mutate by single letter code","M",
lambda: mutate_by_entered_code())

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



#mutate active residue to entered residue code (upper or lower case single-letter)
add_key_binding("Mutate by single letter code","M",
lambda: mutate_by_entered_code())

Many pre-packaged functions available in COOT API. Mostly documented in online manual. Very easy to write your own! Useful e.g. for scripting domain-wise rigid body refinement.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).

The Coot Heer Manual X +			
	• •		-
🖉 🕐 🖬 mttps://www.amrc-mino.cam.ac.uk/verisona/perissie/cool/web/docs/cool.nitmi 📖 C 🔍 Search 🙀 🖬 🔍 🖬 🥀 🖬		8	=
Next: ncs, Previous: filter, Up: Scheme Scripting Functions			
3 11 coot mi			
3.11 cool-gui			
– procedure: run-gtk-pending-events – procedure: coot-gui			
Fire up the coot scripting gui. This function is called from the main C++ code of coot. Not much use if you don't have a gui to type functions in to start with.			
- procedure: handle-smiles-go tlc-entry smiles-entry			
The callback from pressing the Go button in the smiles widget, an interface to run libcheck.			
– procedure: smiles-gui			
smiles GUI			
- procedure: generic-single-entry function-label entry-1-default-text go-button-label handle-go-function			
Generic single entry widget			
Pass the hint labels of the entries and a function that gets called when user hits "Go". The handle-go-function accepts one argument that is the entry text when the pressed.	ie go button is		
- procedure: generic-double-entry label-1 label-2 entry-1-default-text entry-2-default-text check-button-label handle-check-button-function go-button-label handle-	-go-function		
handle-go-function takes 3 arguments, the third of which is the state of the check button.			
if check-button-label not a string, then we don't display (or create, even) the check-button. If it *is* a string, create a check button and add the callback handle-c function which takes as an argument the active-state of the the checkbutton.	heck-button-		
- procedure: generic-multiple-entries-with-check-button entry-info-list check-button-info go-button-label handle-go-function			
generic double entry widget, now with a check button			
OLD:			
	- C-1 - 1 - 1		

Many pre-packaged functions available in COOT API. Mostly documented in online manual.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Lots of key bindings, and easy to define custom keys. Learn them. They make everything much faster.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...

Lots of key bindings, and easy to define custom keys. Learn them. They make everything much faster.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...

You can convert an mrc to mtz using phenix.map_to_structure_factors.

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...



- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...

The optimal sharpening B-factor will vary across the map (assuming some variation in local resolution). On-the-fly adjustment is therefore very useful. (*phenix.auto_sharpen* is good for determining



- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...

- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...



- Simple, intuitive interface for building and manipulating atomic models in density maps.
- Low computational requirements
- Extensive API easy to script or modify (using simple Python code)
- On-the-fly sharpening and low pass filtering (for MTZ).



Successfully read coordinates file /Users/olibclarke/Dropbox/ryr_models_paper2/cam/reprocess/best_tet_resampled_apr12_rs2_real_space_refined.pdb...

- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.





- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.



- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.



- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.



(mol. no: 6) CA /1//-1 ALA occ: 1.00 bf: 50.00 ele: C pos: (251.69,255.83,193.56)

- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.



(mol. no: 6) CA /1//-1 ALA occ: 1.00 bf: 50.00 ele: C pos: (251.69,255.83,193.56)

- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.



- Semi-automated helix placement
- Place cursor at the center of the helix and trigger "Place helix here" (I suggest via a key binding - "h" with coottrimmings)
- Coot will attempt to automatically determine the length and direction of the helix.
- Trim/extend, adjust weights, then refine using real-space refine zone. Drag into density to adjust fit.





- Sequence assignment.
- Adjust numbering to match expected position in sequence.
- Mutate to match sequence
- Fill sidechains manually.
- Adjust sequence register to optimize local fit to sidechain densities.



(mol. no: 6) CA /1//12 ALA occ: 1.00 bf: 50.00 ele: C pos: (241.31,258.23,210.37)

- Sequence assignment.
- Adjust numbering to match expected position in sequence.
- Mutate to match sequence
- Fill sidechains manually.
- Adjust sequence register to optimize local fit to sidechain densities.



(mol. no: 6) CA /1//12 ALA occ: 1.00 bf: 50.00 ele: C pos: (241.31,258.23,210.37)



- Sequence assignment.
- Adjust numbering to match expected position in sequence.
- Mutate to match sequence
- Fill sidechains manually.
- Adjust sequence register to optimize local fit to sidechain densities.



(mol. no: 6) CA /1//12 ALA occ: 1.00 bf: 50.00 ele: C pos: (241.31,258.23,210.37)



- Sequence assignment.
- Adjust numbering to match expected position in sequence.
- Mutate to match sequence
- Fill sidechains manually.
- Adjust sequence register to optimize local fit to sidechain densities.



(mol. no: 6) CA /1//12 ALA occ: 1.00 bf: 50.00 ele: C pos: (241.31,258.23,210.37)



- Sequence assignment.
- Adjust numbering to match expected position in sequence.
- Mutate to match sequence
- Fill sidechains manually.
- Adjust sequence register to optimize local fit to sidechain densities.



(mol. no: 6) CA /1//12 ALA occ: 1.00 bf: 50.00 ele: C pos: (241.31,258.23,210.37)

Use 'Add Terminal residue' to extend chain.

Types of errors in macromolecular models

- Identity (e.g. wrong domain)
- Directionality
- Topology/connectivity
- Register
- Rotamer
- Backbone torsion
- Ligand identification and placement

Types of errors in macromolecular models


Types of errors in macromolecular models

- Identity (e.g. wrong domain)
- Directionality
- Topology/connectivity
- Register
- Rotamer
- Backbone torsion
- Ligand identification and placement



Types of errors in macromolecular models

- Identity (e.g. wrong domain)
- Directionality
- Topology/connectivity
- Register
- Rotamer
- Backbone torsion
- Ligand identification and placement



Strategy for identifying and correcting errors.

- Analyse as you go "sanity checks" on chemistry, nonbonded interactions, surface composition. Use Molprobity for clashes, Chimera or pymol to check e.g. for buried polars, exposed hydrophobics. Monitor agreement with secondary structure, disorder predictions.
- Use EM-ringer to identify errors in backbone and rotamer geometry.
- Look at everything! Manually check and recheck the fit of every residue in Coot. Tedious but necessary.
- Sometimes, you just can't tell the right answer. Don't be afraid to specify sequence ambiguity (use UNKs).
- Half-map FSCs are only really useful to analyse overfitting they tell you nothing about the local quality or correctness of the model.

Model building in EM – RyR1 as an example



- Starting information:
- Crystal structure of first 500aas
- Homology models for four domains (SPRY1-3, RY12, RY34);
- TM fold apparent from density
- Allowed building of polyalanine trace including known and unknown domains
- Fortuitous 3D class with disordered protomer allowed determination of protomer boundaries

Model building in EM – RyR1 as an example



- Starting information:
- Crystal structure of first 500aas
- Homology models for four domains (SPRY1-3, RY12, RY34);
- TM fold apparent from density
- Allowed building of polyalanine trace including known and unknown domains
- Fortuitous 3D class with disordered protomer allowed determination of protomer boundaries

Our initial model allowed location of all previously predicted domains, and several new ones.



Our initial model allowed location of all previously predicted domains, and several new ones.



Improved resolution allows building of a near-complete atomic model. Refined using phenix.real_space_refine



Low local resolution in periphery prevents model completion.



Residues 1250-1650, 2500-3640 still poly-ala – no sequence assignment.

Is lower local resolution in periphery due to unresolved movement of shell?





searches greatly improves local resolution.



Masked refinement with local angular searches greatly improves local resolution.



Before

Allows sequence assignment in regions that previously lacked connectivity and/or sidechain definition.



After

Allows sequence assignment in regions that previously lacked connectivity and/or sidechain definition.

Thanks for listening!