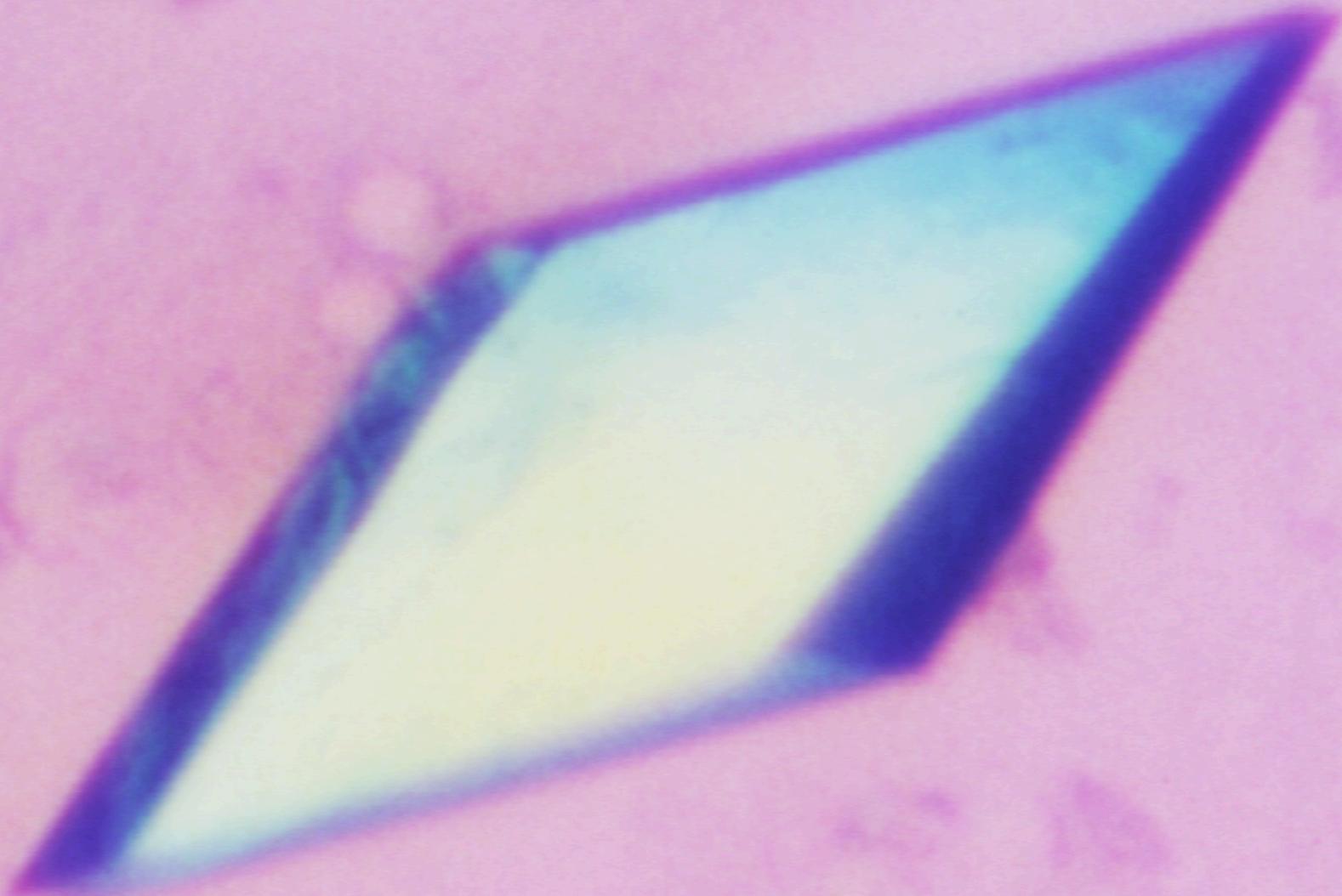
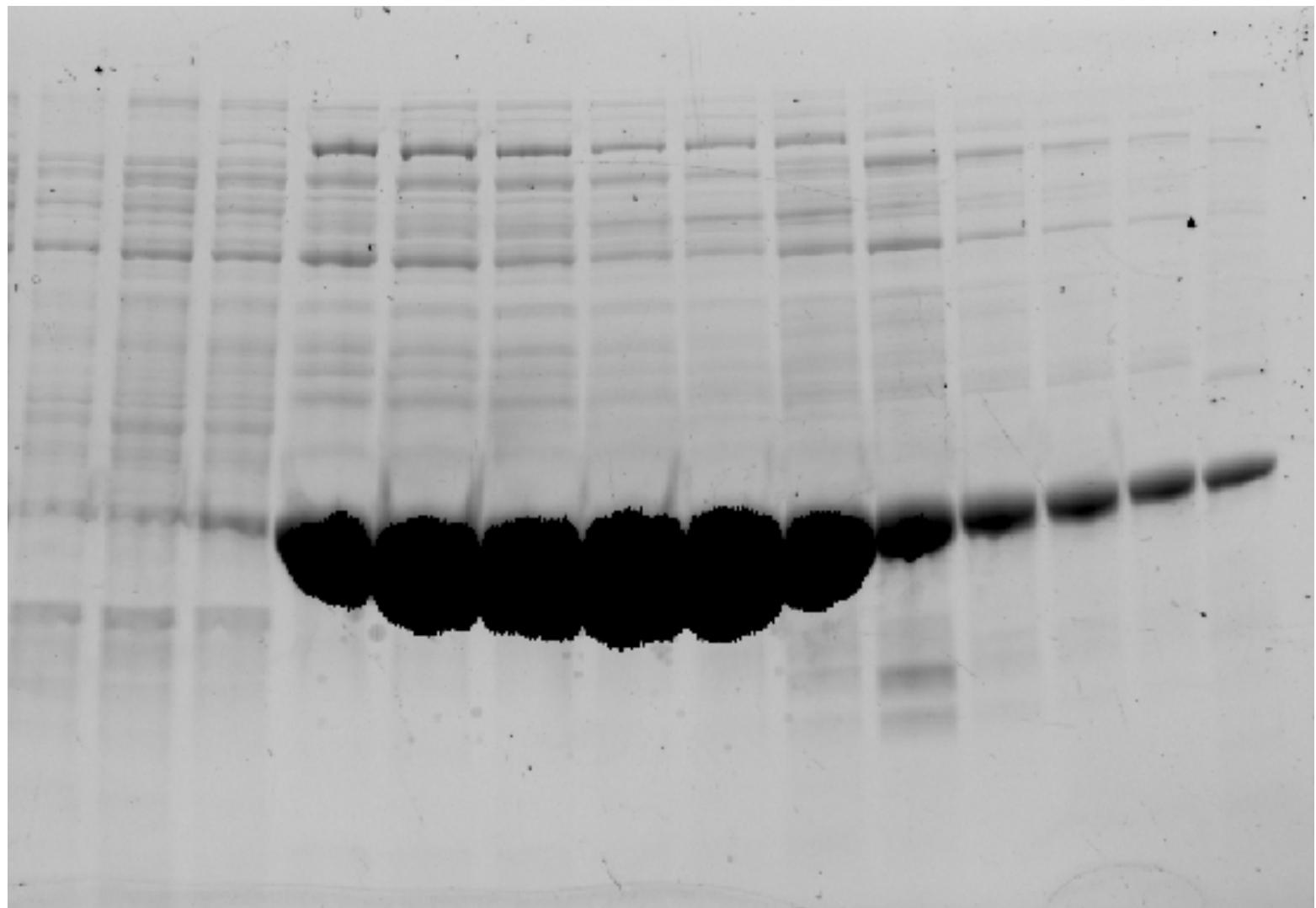


Expression, purification, characterization, and crystallization

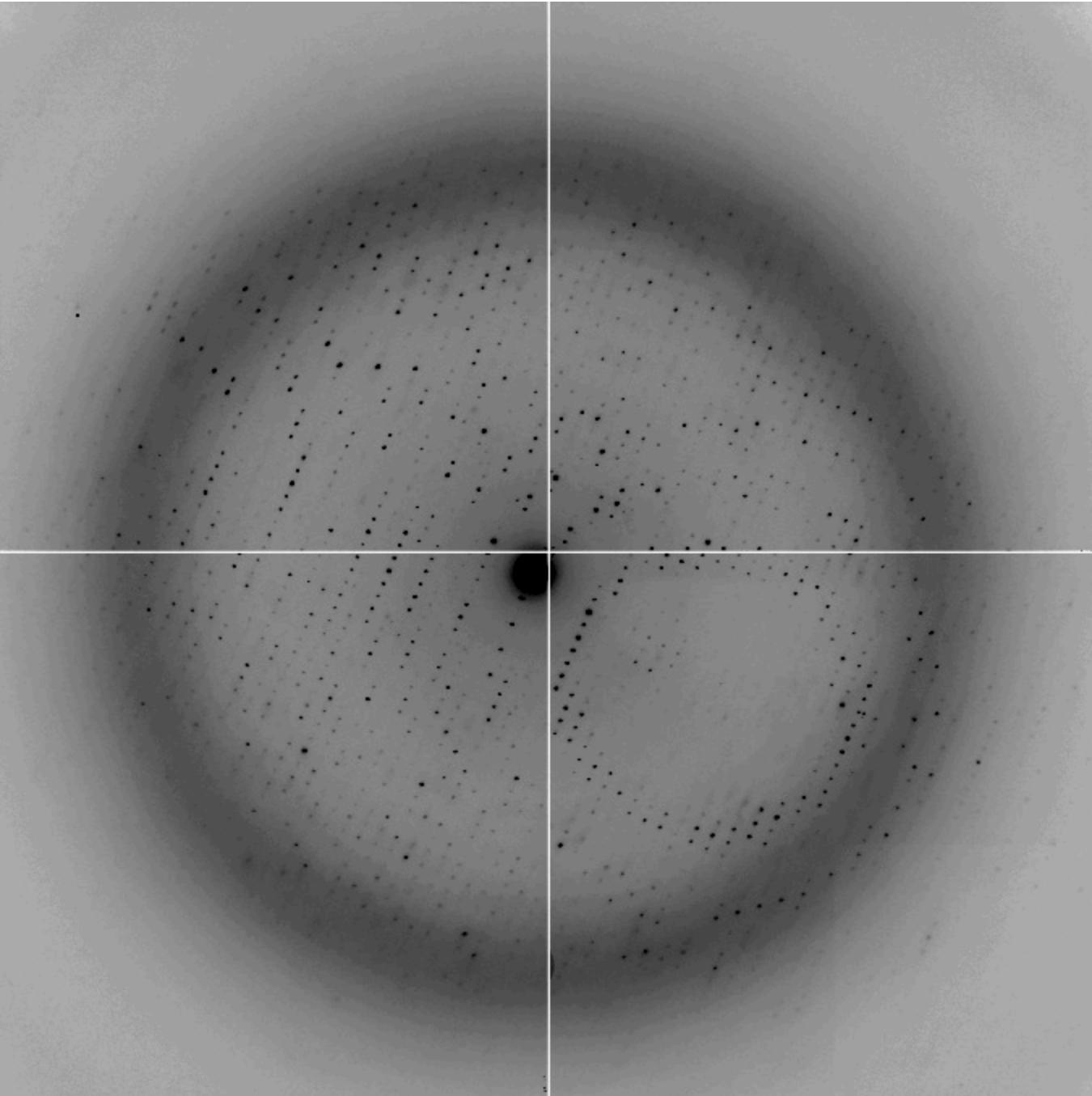








© Carsten Peter/National Geographic/Speleoresearch



So how does one go about solving a crystal structure?

formulate question	very hard!	Boss
make sample	cloning, expression, purification	you
make crystal	screening, optimisation	you
collect diffraction data	synchrotron, integration, scaling	Post-doc
solve phase problem	MR, SIRAS, MIRAS, SAD, MAD, hybrid	Post-doc/Randy/Phil
build model	manual or autotracing	Post-doc/you
refine model	agreement of model and data	Post-doc/Garib
interpret model	very hard! back to top?	Boss

... and might fail at any step!

This is the part where YOU make all the difference!!

Disclaimer:

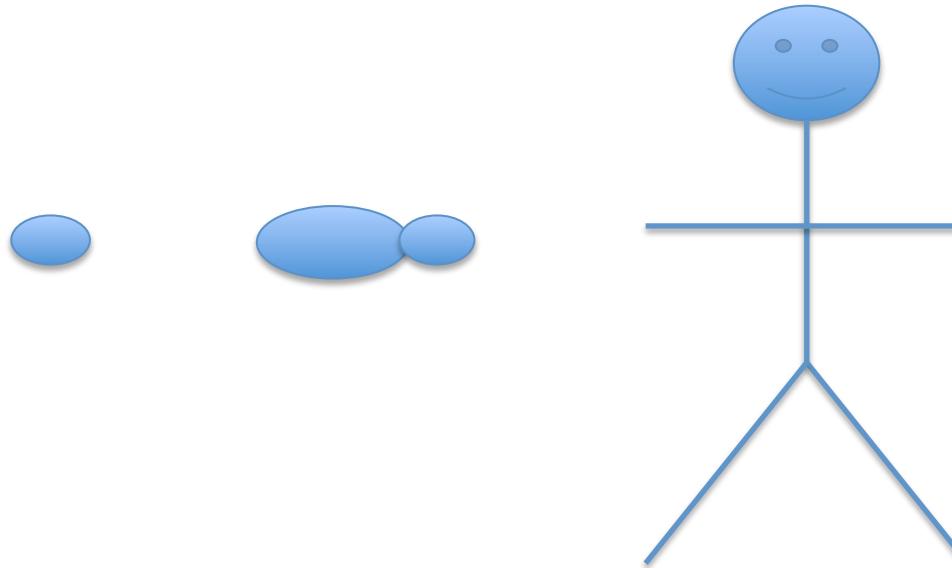
This is a limited (i.e. my) view of how to do crystallography

You have to find out what works for you

(But some of this might be useful)

What to crystallize: Different species

- Bacterial: Great diversity, Easy to work with
- Archeal: Half way between bacterial and eukaryotic
- Eukaryotic: more difficult, but sometimes essential



What to crystallize: Different species

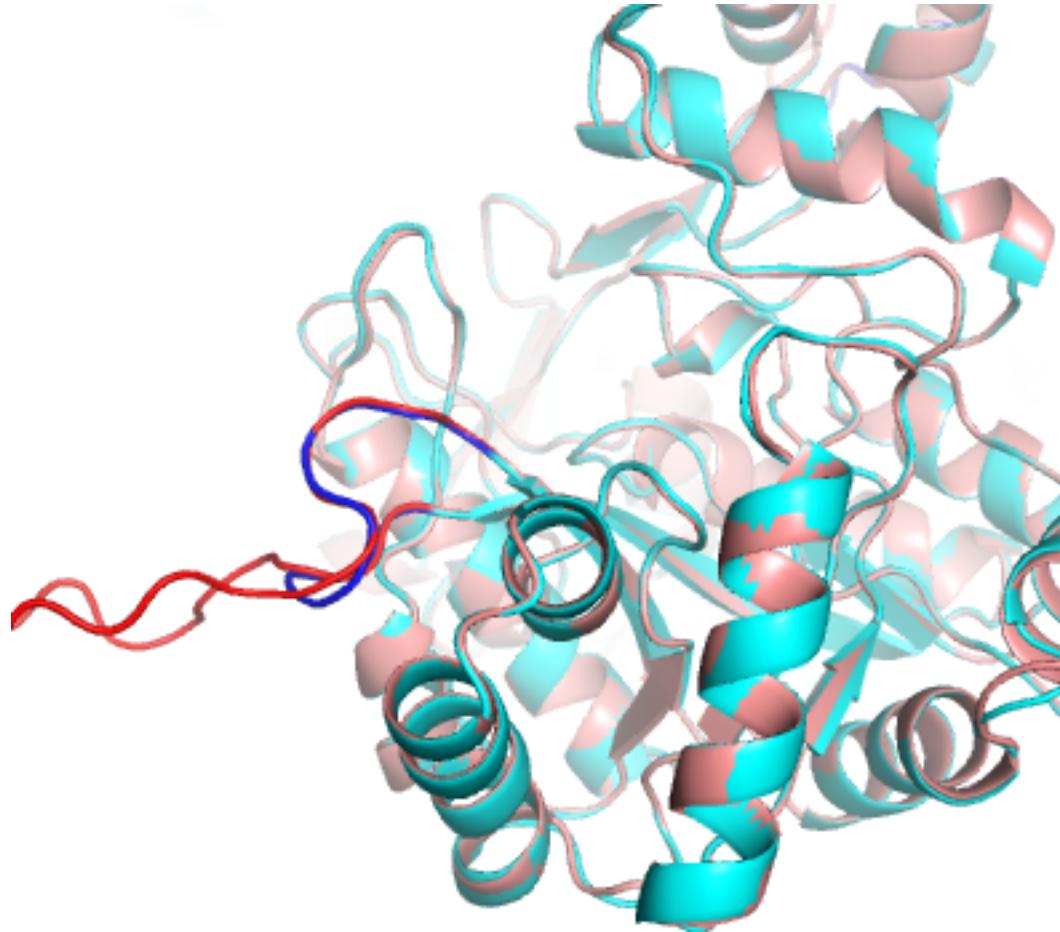
ecoli_dp3a MSEPRFVHLRVHSYDSIMDGLAKTAPLVKKRAALGMPALAITDFNLCLGLVKFYGAGHGAGIKPIVGADPNVQCDLILGDELTHLTVLAANNTGYQNLTLLISKATQRGTYC-AAGPIIIRDMLIELNEGLILLSGGRMGDVGRSLLRGNSA
 salty_dp3a MSEPRFVHLRVHSYDSIMDGLAKTAPLVKKRAALGMPALAITDFNLCLGLVKFYGAGHGAGIKPIVGADPNVQCDLILGDELTHLTVLAANNTGYQNLTLLISKATQRGTYC-AAGPIIIRDMLIELNEGLILLSGGRMGDVGRSLLRGNSA
 vibch_dp3a MSDPDKFIHLRIHSDFMSVMDGLSKVPLVKKVAAMGMPALMTDFNLCLGLVKFYTAHNCGVKPIIGADFTLQSEEFDETLKLTLAKNNVGYKNLTLISKAYLRGHV-OHQPVIDTKRALWVEHAEGLIVLSSGGSXGEVGRALLKGNOQ
 haein_dp3a -SQPRFPIELRTHIDFSMIDSIVKVPLVKAACAANEMVAMGLTDFTNFCCGVVFYSEMLSSGMKPPIIGADVKVKSPLCGDEYFDLTLAKNNEGYRNTLILSKATQRGTYC-NDLPYIDQDWLIEERDGVIILSGGTQGDGVGKLLKGHAA
 pasmu_dp3a MSEPRFVHLRVHSYDSIMGIAVVKPLVKAACVDFNVMHVGMLTDFTNFCCGVVFYSEMLSSGMKPPIIGADVKVKSPLCGDEYFDLTLAKNNEGYRNTLILSKATQRGTYC-QDLPLYDQMAAHLEREGIIVLSSGGTQGDGVGKLLKGHAA
 bucai_dp3a MNEPRFVHLRVHSYDSIMDGLSKPDLVKKASLNPAAITDFNLCLGLVKFYTAHNCGVKPIIGADVLVRSLELCGDEPFELTLLAKNNIGYHNTLILSKATQRGTYC-NDLPYIDQDWLIEERDGVIILSGGTQGDGVGKLLKGHAA
 bucpar_dp3a MNLEEFVHLVHSYDSIVDGLSKPEDLILKSVNLMGKALSITDFNLCLGLVKFYTAHNCGVKPIIGADVLVRSLELCGDEPFELTLLAKNNIGYHNTLILSKATQRGTYC-DNLPLYDQMAAHLEREGIIVLSSGGTQGDGVGKVLNCQSS
 pseae_dp3a ---SFVRLRLHTEFSLVGLVRLVPLAKAVAGLGMPPAVVTDQSNCMSLVRFKYTAHMAGIKPIGADINLASREEDGPLSRSLSLAMNAKGRYNTLTELISRGMWSQGR-NGEIIIERDWDVKEAEGLIALSAKEGEIGHALLDGEAA
 psefl_dp3a ---PASFVRLRLHTEYSLVGLVRLVPLKLVALTGMGMPAVVTDQSNCMSLVRFKYTAHMAGIKPIGADINLSRNKDPDGLPSRSLSLAMNAKGRYNTLTELISRGMWSQGR-NGMVIIEREWAEGLIALSAKEGEIGHALLDGEAA
 xylyf_dp3a MSTSSFVHLHTEFSLADSTAKQANLISRAVELGLPLAVTDNLNLPALIHKFYKAETVGKIPBQGADLLIA-EPEHPPGMGTLICRDHAGYLNLQLSISRGWLEGRHPEGGVAHVPDFWRDHKNLFALIG-RHSLAGOLLAGKGRAD
 xylifa_dp3a MSTSSFVHLHTEFSLADSTAKQANLISRAVELGLPLAVTDNLNLPALIHKFYKAETVGKIPBQGADLLIA-EPEHPPGMGTLICRDHAGYLNLQLSISRGWLEGRHPEGGVAHVPDFWRDHKNLFALIG-RHSLAGOLLAGKGRAD
 neimb_dp3a MTEPTYIPRLRHTEFSITDGMVRIRKKLIAKQEIICLPLAGISDLNMFGLVFKYACRSAGIKPIGADVRIGNPDAFDPKPFRMLIIRNDAGYLNLQLSELLTRAYGKDRN-VHHABLNPEWLENDNSGLICLSGAHYGEVGVNLNGED
 neima_dp3a ---WIPLCHQSQVSILDATCSIKKFVAKAVEYQIIPALALTDHGNLFGADEVFYKTCQNAIKPIICCCELYVAPSSRFDKANHLILCKDEEGYRNLCLLSSLAYTEGFYVV-PRIDRDLLSQHSKGKLGICLSCASLGSVAQAAL-SEE
 chltr_dp3a ---WIPLCHQSQVSILDATCSIKKFVAKAVEYHIPALALTDHGNLFGADEVFYKTCQNAIKPIICCCEFDKKERKSQVANHLLCDEEGYRNLCLLSSLAYTEGFYVV-PRIDRDLLSQHSKGKLGICLSCASLGSVAQAAL-SEE
 chlmu_dp3a ---WIPLCHQSQVSILDATCSIKKFVAKAVEYHIPALALTDHGNLFGADEVFYKTCQNAIKPIICCCEFDKKERKSQVANHLLCDEEGYRNLCLLSSLAYTEGFYVV-PRIDRDLLSQHSKGKLGICLSCASLGSVAQAAL-SEE
 chlpn_dp3a ---WIPLCHQSQVSILDATCSIKKFVAKAVEYHIPALALTDHGNLFGADEVFYKTCQGIPIICCCEFDKKERKSRAAHHLILCKNEQGYRNLCLLTSALTEGFYVV-PRIDKDLRQYSEGLICLSCASLGSVAQAAL-SEE
 helpj_dp3a -ENKAFTLHLHTEYSLLDGANKIKILAKRIKELGKMSVSVDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRINKLRLREHSKGIIAASSACLOGEVNVHLNTNNED
 helpy_dp3a -ENKAFTLHLHTEYSLLDGANKIKILAKRVKELGKMSVSVDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRINKLRLREHSKGIIAASSACLOGEVNVHLNTNNED
 canjje_dp3a ---QFTBLHLHTEYSLLDGANKIKLSELALIHKEQGATSVAMTDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRINKLRLREHSKGIIAASSACLOGEVNVHLNTNNED
 borbu_dp3a ---RFIHLHVHSYDSLLDGAAKISDIISKAKKCNMSHIALTDHGNLFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRIDKDDLEKYSSEGISTSACIGGLIPRLLRILANRFE
 theaq_dp3a ---SKLKFAELHQHTQFSLLDGAAKIQDLDLKTETPEDPALATMDHGNLFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRIDRDLRQYSEGLICLSCASLGSVAQAAL-SEE
 deira_dp3a ---OPKFKFAELHQHTQFSLLDGAAKILKDLWKAKEVOTPALATMDHGNMFGADEVFYKTCQNAIKPIICCCEFDKKERKSQVANHLLCDEEGYRNLCLLSSLAYTEGFYVV-PRIDKDLRQYSEGLICLSCASLGSVAQAAL-SEE
 myctu_dp3a -ACSSFVHLNHTEYSMLDGAAKITPMLAEVERLGMPAVGMDTHGNMFGAIDPYTSNKAICLAKGIKPIIGVFDTRRILNCDSYTHLMMENATGLRNLFKLSSHASFEQQLSKN-SRMDAELIAEHAEGIITTGCPSEGVQTRL-LQD
 mycyle_dp3a MNQSSFVHLNHTEYSMLDGAAKITPMLAEVERLGMPAVGMDTHGNMFGAIDPYTSNKAICLAKGIKPIIGVMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRINKLRLREHSKGIIAASSACLOGEVNVHLNTNNED
 strco_dp3a MSKPPFTLHLVHTQVSLLDGAAKILKDMFDACNEMGMSHIAIMSDDHGNLHLGAIDFPHSAKAGVTPPIIGEAYAWESPRRNKYTHKTMATNSKLHNLVLSLDAAEYGJLQKN-PRMDKETISQNSWEGIVASTGCPSEGVQTRLRHLGHD
 trepa_dp3a MARMSFVHLHVHSNYSLLDGASSLQLRVLRTAKSLGQEAALTDHGNMFGALHFQKVCVSAEGIKAIICCEPDRSEHTIGRRYYHLIVLAKNETGYRNLMLVLSKAYIEGMYK-PRVDDELLAQRAEGLICLSSCLAGQLPYLLLQRKR
 aquae_dp3a ---FVHLHLHTEFSLDGAAKIKDELVLKKAKEYVGKAVGMSDHGNLFGAIDPYTSNKKBEGIKPIIGMEFDRKTTSBEDNITHLIAKDDGKLNLKML5TLAYKEGFYVV-PRIDHELLQEEHKKJIAFSGCLGEVQFQILMQRRED
 staaau_dp3a -YLNIAHTAYDLNSSLKIEDAVRLAVSENVDALA1DTDNVLYGPFKFYDACIANNIKPIFGMTIVVENGLY-TVETVVLVAKNNNDGLKLQSLSSIEKMMNAEH-VSFELLKRFSNMIIIFKVKV-D
 staaw_dp3a -YLNIAHTAYDLNSSLKIEDAVRLAVSENVDALA1DTDNVLYGPFKFYDACIANNIKPIFGMTIVVENGLY-TVETVVLVAKNNNDGLKLQSLSSIEKMMNAEH-VSFELLKRFSNMIIIFKVKV-D
 staam_dp3a -YLNIAHTAYDLNSSLKIEDAVRLAVSENVDALA1DTDNVLYGPFKFYDACIANNIKPIFGMTIVVENGLY-TVETVVLVAKNNNDGLKLQSLSSIEKMMNAEH-VSFELLKRFSNMIIIFKVKV-D
 staep_dp3a -HLNITSHFDLSSSLRIDALIKKACRKYCAEFSALALTDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRFVSEFLLQFSSNLLIIIFKVKV-D
 bachd_dp3a -VHIVHSEVTEYLSSMCRIPALVEKAKAAQFSALALTDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRFVSEFLLQFSSNLLIIIFKVKV-D
 bacsu_dp3a -SFVPLQVHGSYSLNSAAAEEVLVBSEADRGLYASALATDDHGNMFGAIDPYTSNKKBEGIKPIIGMEAIHNDKETKQRFLCLFAKNQEGYENMLPLSSMAYLEGFYVV-PRFVSEFLLQFSSNLLIIIFKVKV-D
 ricpr_dp3a -QPEFIRLRTQSSYFSLESALTIKVEVLLALHKMPALTSDRGNLFGSLEFLSYRKKKKLPIHGVLINQYDIN-AFAQILLIAKDETGYKNNLKLSSLTFTKNSDR-KICEHIGFEDLKYQEGVIALCCYTGDIVGKCLLARKQE
 strp8_dp3a -FAQLDTKTVFSFMDSLIDLNHYFERAKQFGYHTIGIMDKDNLGYAIFIKGQCKNGLQPVLGLEVEILYQER-QVLLNLIAQNTQGYHQLLKLIST-AMMSGKL-HMDYLCQHLEGIAVIIIPSKG
 strp3_dp3a -FAQLDTKTVFSFMDSLIDLNHYFERAKQFGYHTIGIMDKDNLGYAIFIKGQCKNGLQPVLGLEVEILYQER-QVLLNLIAQNTQGYHQLLKLIST-AMMSGKL-HMDYLCQHLEGIAVIIIPSKG
 strpy_dp3a -FAQLDTKTVFSFMDSLIDLNHYFERAKQFGYHTIGIMDKDNLGYAIFIKGQCKNGLQPVLGLEVEILYQER-QVLLNLIAQNTQGYHQLLKLIST-AMMSGKL-HMDYLCQHLEGIAVIIIPSKG
 lacla_dp3a -FAPLNTKTEYESFLDSVVKVUDYLYTAHRLQYGTQVYCICDVGVLNHAFAFRVRAQKFNLPIIISIENFWERGL-PIAFSFIADKTEGYKNNLKLIST-LHNYGRR-QFSDIQNHLSSGIALIIPETYGS
 urepa_dp3a -FINLNVHSYSLNLSLIDDLLKVALDKQFVYVLTDDLNMMYGCJEFYDRAKAHNLTPIIGLEFEYQN-ATLVAFAKNYGYLKLKIKNSSWIMTSKED-BIQDDFDLIIIVCK-KG
 yerpe_dp3a -MKALMVRIDTSLGESALKAENAVIARDAGTAVISADMNSIASVPLQRAAG-DDMAVICGVKLNUVDDLVRDRSICFTALIKEQF-FVPRALDQDAAKGNIIILS DIGSVPQR-R
 ruler 1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120.....130.....140.....150

What to crystallize: Different constructs

	Construct A				Construct B				Construct C				Construct D			
ecoli_dp3a	DTKKLNRRVLEKLIMSGAFDRLGPHRAALMNSLGDAKADHQAKAAE--	IGADMFGVLAEEP-	EQIEOSIASCOPWPEQVVLGDGERETLGLYLTHGPING-	LKEIERVVG--	VRLKDMHP	TERGVITIAAGLVVAARVMVTKRG-NRIGI	147									
salty_dp3a	DTKKLNRRVLEKLIMSGAFDRLGPHRAALMNSLGDAKADHQAKAAE--	IGADMFGVLAEEP-	EQIEOSIASCOPWPEQVVLGDGERETLGLYLTHGPING-	LKEIERVVG--	VRLKDMHP	TERGVITIAAGLVVAARVMVTKRG-NRIGI	147									
vibch_dp3a	DLKKVNKRVIKEKLILLAGALDRLGPHRAAMMASVDDAVRAASQHQAEAA--	FGQADMFGVLEEDAP-	EEVEQKYEQQPPEWPEKVRLEGGERETLGLYLTHGPVDE-	YLKEELTKYTS-	CRLNEAAP-	TRRDQSLIVAGLVIARVMVTKRG-TRIGL	146									
haein_dp3a	DLKKINRRTFESLILSGAFDKLGPGRHAALESKNLEDALRASDQHAKDEA--	MGQADMFGVLEEDSH-	EDVENAATNPPTYEKQILDGERETLGLYLSSHPVSR-	YLKEELSHYT-	TRLKDLAP-	NRRGQISTVAGLIVVAARIAMTKKG-NRIGI	146									
pasmu_dp3a	DLKKINRRTFESLILSGAFDKLGPGRHAALESKNLEDALRASDQHAKDEA--	MGQADMFGVLEEDSH-	EDVERAIASTPRNSEEKYLEGERETLGLYLSSHPISP-	YLKEELANYS-	TRLKDLVP-	NSRGMMSIVBGLLVSRRPAVTKKG-NRIGI	146									
pseae_dp3a	DLKRINKRTTLEALIRAGALDRLGPHRAVLLAAMEEAIQAB-QTARSHD--	SGHMDLFGVFAEPE-	ADVIAHHRKVKEELKLKERLKGEKDITLGLYLTHGPIDE-	VEGEVRRFAR-	Q-VELKP-	AR-DTQTVAGLIVNLRVVMKNGG-DKMGF	142									
psefl_dp3a	DLKRINKRTTLDGLIRSGALDRLGPNRAVLLAAMEEAIQAB-QTARSHD--	SGHADLFGFLVVED-	ADVIAHHRKAKEELKLKERLKGEKDITLGLYLTHGPIDE-	VEGEVRRFAR-	QRIIDLKP-	AR-DTQTVAGHIIALRVVMKNGG-DKMGF	144									
bucai_dp3a	DSQRXITRRVLEKLIMSGSCDFDKNRNYLLOSIDDAINASKESFRIKS-	FKDQSLFGIFQNEL-	NOVKNNNLVNLVCPEKNNLQNEYQVLFYLTSHGPING-	YKKELEYCVNG-	VRLSQQLK-	IKRNKKLILVAGIIVSIIKIKITKKG-NRIAI	147									
bucap_dp3a	DPNKTRKVLVEKLIMSGSCDFDKNRNYLLOSIDDAINASKESFRIKS-	FQDESLFGFKEEL-	NILKNNNLVNLVCPEKNNLQNEYQVLFYLTSHGPING-	YKKELEYCVNG-	VRLSQQLK-	FNHHKKIMVAGIVVSIIKIKITKKG-NRIAI	147									
xylft_dp3a	TSALKLNRRALEAMIRHAGALDELGKRNRAVSMQLPEVIRAKEMSRERE-	SGCNSLFGNADPG-	PVILQDLPECEEENPLTRMLNGERETLGLYLTSGHGPFD-	TRKQVKELVG-	LGSQQQRNGEKRWTQPVENNTILAGLVLVSVR--	RKG-DSQVF	147									
xylifa_dp3a	ASAALKLNRRALEAMIRHAGALDELGKRNRAVSMQLPEVIRAKEMSRERE-	SGCNPNSLFGNADPG-	PAIQILQDLPECEEENPLTRMLNGERETLGLYLTSGHGPFD-	TRKQVKELVG-	LGSQQQRNGEKRWTQPVENNTILAGLVLVSVR--	RKG-DSQVF	147									
neimb_dp3a	GKEHMNRRTTLEALIRGGAFDSIEPNRAMLLANIDLAMNNADQKAAN--	ANGQGLFDMMEDAI-	EPVR-LIDAPMWNSSEKLEAEEKTVIGFLYLSGHGPFGP-	YAOEVROIAPI-	TLDLR-	LKPQDSVRLAGFVTAVRTMM-GKR-GKIAF	138									
neima_dp3a	GKEHMNRRTTLEALIRGGAFDSIEPNRAMLLANIDLAMNNADQKAAN--	ANGQGLFDMMEDAI-	EPVR-LIDAPMWNSSEKLEAEEKTVIGFLYLSGHGPFGP-	YAOEVROIAPI-	QLSK-	LKPQDSVRLAGFVTAVRTMM-GKR-GKIAF	138									
borbu_dp3a	DDKVHNKKFLESIAKSGLFDSDQNLKTLFENHLIEVVSEDKNNNK-	LGQNSLFGALSQDP-	IQQSFTNYQTFKYESSELLGFEEKLLGFYVSGHPLD-	YKKAIDSFSS-	LNVLTDLA-	AKKDSIVQFSGILNSVVKIQTCKRN-NKMAF	146									
trepa_dp3a	PATSLNNKNAEIMIXAKSGCDFRFGVTRASLJAHLDAMKTYVARKKAVTS-	SRQASLFDALTDLG-	ECSEYTPVMEEENSRERLRIEKELMGYTSIISGHPLDE-	YRSVIGEKAT-	LDLGHIEEN-	ARSENKYLIVGVVLNAIHPYITTKKG-KRMAF	145									
theaq_dp3a	PEQVUNNKRALESLVKAGALDAFG-DRARLLASLEPLLRMMAAESTR-	GRSGLVLGFAEVSE-	-PPLVVEASP--LDEITMLRYEKEALGIYVSGHGPVLR-	YLREVASCIIIELSEFVRELPG-	K-PKVLLSGMVVEVRKPFTRSG-GMMAR	140										
deira_dp3a	GNVKCNRKRALESLVKAGALDAFG-ERROLIESLEDAAGIAEINARAOGSMSMMFGMEEVVKKE-	RPPLSIA-	YSDLERLAIKEEALGLYISGHGPLEQEGLREAAACRV-	LQNVAPG-	KRKQAVLAGMIEGVVKKPTKG-GMMAR	145										
myctu_dp3a	DISACNKKVTTESLIKAGAFDSLGLHARKGLFLVHSDAVDSVLGCTKAE-	ALGQFDLFGFS-	-DDCAGTAGDPVTIIXVP--WEDKHLALAREMGLYVSGHPLNG-	VAHLLAAQVD-	AIPAIALLD-	VPNDAQVVRVGGILASVNRVRVNKG-MPWAS	147									
mycle_dp3a	DITSCNKKVTTESLIKAGAFDSLGLHARKGLFLVHSDAVDSVLGCTKAE-	ALGQFDLFGFS-	-DDCAGTAGDPVTIIXVP--WEDKHLALAREMGLYVSGHPLNG-	VAHLLAAQVD-	AIPAIALLD-	VSNDIQVVRVGGILASVNRVRVNKG-MPWAS	147									
strco_dp3a	EAVVCKNRKTTESLIKAGAFDEMGTTRKGLQAEYEPMDINVVAKRKE-	AEGQFDLFGGMD-	SDEPDVVFVGEDE--MDKLYLQAQEREMGLYVSDHPLFG-	LEHVLDD-KADA-	GISQLTGGD-	FPGDGAVVITGGIISGLQRKMKQG-NHAWAI	147									
helpj_dp3a	DFSKLTKKSLEPLVKSGLSNDLNLGYTCKTMLANLDAGRAKDNKEMMO-	GGNSLFGAMEGGIK-	-EVVLDLMDVMD-ERHDATCLLECEYETLGIHVSGNPLD-	FKEEIKGFKN-	LVKSIDIIE-	LEIGSQAALLGKIMEVKKKJIKGKG-KPGVPT	147									
helpy_dp3a	DFSKLTKKSLEPLVKSGLSNDLNLGYTCKTMLANLDAGRAKDNKEMMO-	GGNSLFGAMEGGIK-	-EVVLDLMDVMD-ERHDATCLLECEYETLGIHVSGNPLD-	FKEEIKGFKN-	LVKSIDIIE-	LEIGSQAALLGKIMEVKKKJIKGKG-KPGVPT	147									
camje_dp3a	DPTKINRRTLESLIKAGAFDEFGFTTRKALFDNMENLSEASRKMVAEVKRN-	AASSLSFGEEELTSG-	-VQVNFTPKNE- EFEVMEKLYKEEILGIVVSGHPLDR-	FYEQIN-AID-	YVKSIDFES-	LKNNGEILSIGKIEDFKSMSMSKNN-KRIGR	146									
aquae_dp3a	KNRMKINKKVVEALVNLGFTDCKKRNKELLAKVANS-	ERKALMATQ-	-EVVLDLPLKEEVLKGFLYISGHPLD-	YKLLK--MR-	YTPDLEE-	WDKESEAVLIGVITEIJKVVKTKKG-DTMVA	125									
chltr_dp3a	DFKVKTCKQLESNLVDAFTFCFEPNKBDLALAINLDYTFSKREKKEAA-	GVLTFPSLNMARD-	-VKITVSPENVIQRSFKEELLKREKEELLGVVYLAHPMDA-	VEHMLP-FLS-	VPPVAPDFEG-	LPHGIIIRFVFL-LPHGIIIRFVFL-AQNSAEQKKFAL	146									
chlmu_dp3a	DFKVKTCKQLESNLVDAFTFCFEPNKBDLALAINLDYTFSKREKKEAA-	GVLTFPSLNMARD-	-VKITVSPENVIQRSFKEELLKREKEELLGVVYLAHPMDA-	VEHMLP-FLS-	VVQSKDFEEG-	LPHGGSVVRUVFL-LPHGGSVVRUVFL-KVTKTISSEVKHFAL	146									
chlpn_dp3a	DLKVKSSKIESLIDARGCFDCFDNSRDLILLASPEVLEIADKKEEAA-	GVMFTFFGAMDRK-	-EVPICLPKDIPTRSKKELKKERELLGIYLTTEHPMDT--	VRDHLRS-RLS-	VVLAGEFEN-	LPHGGSVVRUVFL-LPHGGSVVRUVFL-KVTKTISSEVKHFAL	147									
ricpr_dp3a	PPKINSKLLNENLIKAGGFDELHNDNRLQLFLSIPKLIAYSTSYHQB-	-ESNQFSLIKVS-	-LPTILVSSDYADRNLTAFYIPEAMGLFLSNHPLTEYQGISRLLNTR-	DLYNKLPS-	GHNRVVLAGVIOKKDSRMSARG--RFVT-	142										
yerpe_dp3a	EKK-ANSRVRRESLQNVGAFASIE--PGSLPATDPERLRLQDQALMG-	-LVIDAVKASRPFEMPKRS-	--AEVNVLMT-	-RMAAEMLGLDDBLIRP-SIGIKPKIMTG	-RMAAEMLGLDDBLIRP-SIGIKPKIMTG	96										
bacsu_dp3a	PSKSVNRMKLEALIFSGAMDEFCONTRLLASIDVALEHAELFAADD-	-QMLGLFDESFSIK-	-P-KIVETSELPLVLDLAFEKEETLGIYTSNHPLSA-	-FRKQQLTAQQA-	-VSLILQQA-	-VSLILQQA-	140									
bachd_dp3a	PDKIVTSRVMESLIKAGALDELG-ERATLLANIEEAFQFAEVQKEFQE-	-N2GGLFQLSVE-B-	-P-EVIKVPLDDELEKLAYPEKEAVGFYLSGHPLLA-	-VTESLRQYD-	-LTYLEG-	-ERRFVPLLAGMIIRRIRTKRG-VMGF	136									
staau_dp3a	PKRVKTRKLLLEALILVGAFDAFGKTRSTLLQAIQDQVLDBLNLIEQD-	-GFLFDILIP-X-	-Q-MYEDKEELPDALISQYEKEYLGFLYVSQHP-	-VDKKFVAKQY-	-LTIIFKLSN-	-AQNYKPLVQFDKVKQIRTKKG-NMAF	131									
staaw_dp3a	PKRVKTRKLLLEALILVGAFDAFGKTRSTLLQAIQDQVLDBLNLIEQD-	-GFLFDILIP-X-	-Q-MYEDKEELPDALISQYEKEYLGFLYVSQHP-	-VDKKFVAKQY-	-LTIIFKLSN-	-AQNNKPLVQFDKVKQIRTKKG-NMAF	131									
staam_dp3a	PKRVKTRKLLLEALILVGAFDAFGKTRSTLLQAIQDQVLDBLNLIEQD-	-GFLFDILIP-X-	-Q-MYEDKEELPDALISQYEKEYLGFLYVSQHP-	-VDKKFVAKQY-	-LTIIFKLSN-	-AQNNKPLVQFDKVKQIRTKKG-NMAF	131									
staep_dp3a	PKRVKTRKLLLEALILVGAFDAFGKTRSTLLQAIQDQVLDBLNLIEQD-	-EMLFIDILIP-X-	-Q-SYEKEELPDQSLDYEKEYLGFLYISKIP-	-VEKKFEEKK-	-LGIFQLSN-	-GSHVQPLVFDIQCQIRTKKG-NMAF	131									
strp8_dp3a	PEKIQKKVFLPLEPLINKIGLFYFEPNRRKILNDLG--LLVVFVNELG-	-SDSS-	-F-SWVDTKDYSATEKSYSLQEIVGVGMSKHPPLID-	-IAEKSTQTF-	-PISQLVK-	-ESEAUVLQIDSIRIIRTKSGQ-NMAF	124									
strpy_dp3a	PEKIQKKVFLPLEPLINKIGLFDFCPERNRKILNDLG--LLVVFVNELG-	-LPS-	-F-SWVDTKDYSATEKSYSLQEIVGVGMSKHPPLID-	-IAEKSTQTF-	-PISQLVK-	-ESEAUVLQIDSIRIIRTKSGQ-NMAF	127									
strp3_dp3a	PEKIQKKVFLPLEPLINKIGLFDFCPERNRKILNDLG--LLVVFVNELG-	-SDSS-	-F-SWVDTKDYSATEKSYSLQEIVGVGMSKHPPLID-	-IAEKSTQTF-	-PISQLVK-	-ESEAUVLQIDSIRIIRTKSGQ-NMAF	124									
lacla_dp3a	PNNFHKKENILPLIYIGAFDYADSNRGLKAXLNDHALNLLNYYSDIF--M-	-ASSGGG-	-F-AIHEADEYSELEKSYDFEKNLLGIGVPHPLQ-	-LARRPEGNFT-	-PLAQVK-	-NRRMILVBINIYIRTHTRKKG-NMAF	131									
urepa_dp3a	KKGKVSXKNIIEILIRVGTDSFGINRLFLNNNL-	-E-	-I-FEKT-	-GLNGHFFDLH-	-LGV-	-LGDYAKDMSVNR-	66									
ruler	1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120.....130.....140.....150.....160.....															

Max Planck Institute Teubingen toolkit => sequence searches, 2ndary structure prediction
 Clustal (alignments) Jalview (view/edit)

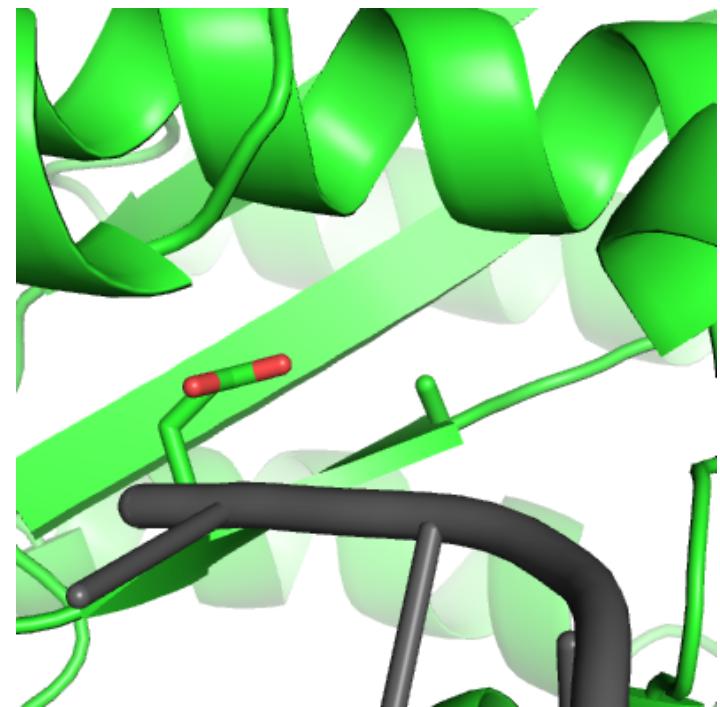
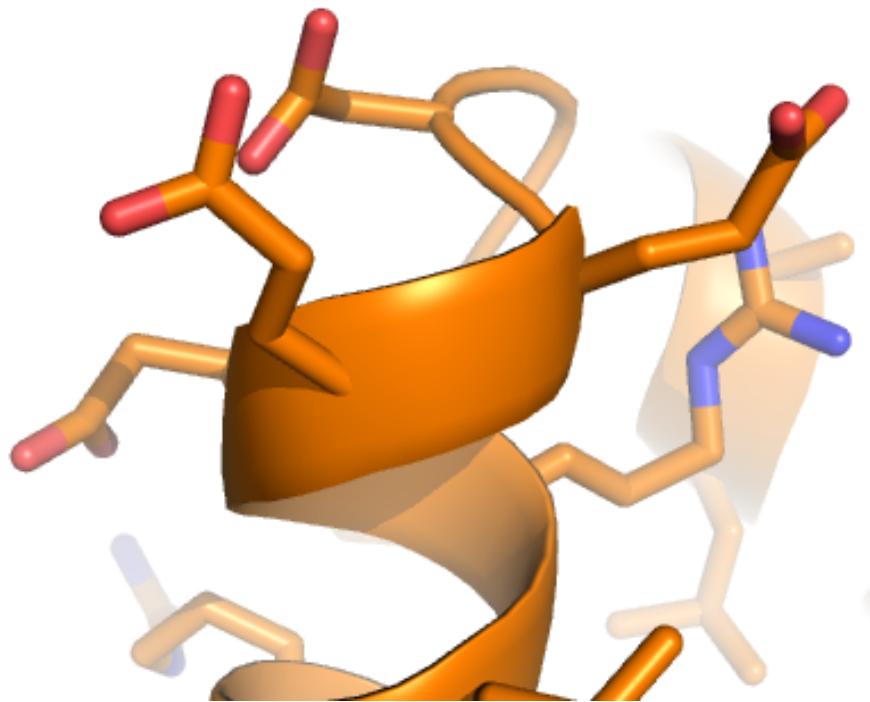
What to crystallize: Different constructs



Modeller (Andrej Sali, UCSF)

What to crystallize: make mutants

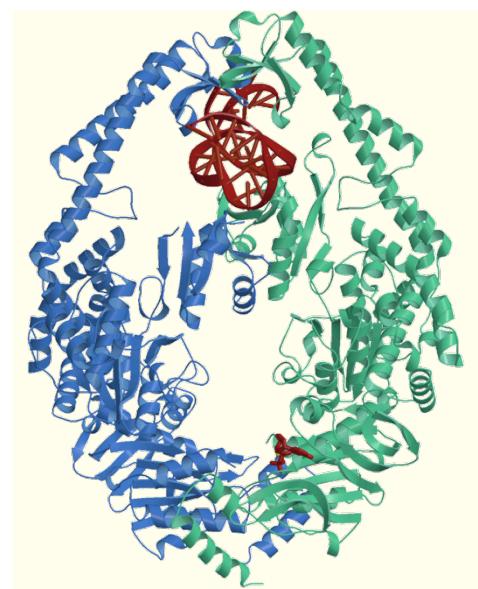
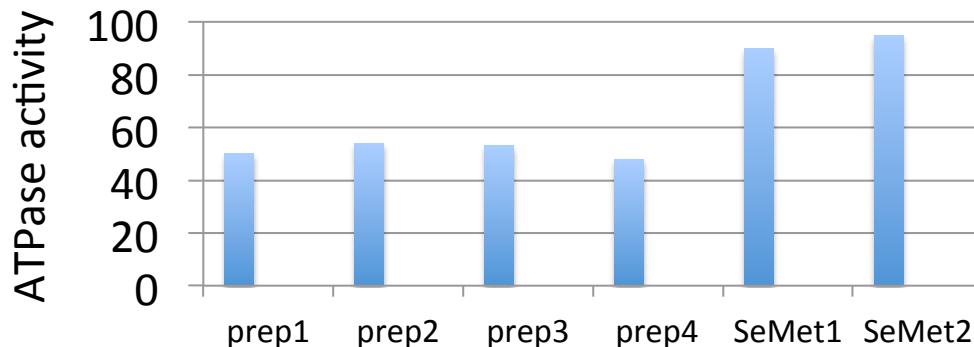
Surface engineering: hydrophylic residues at loops into hydrophobic residues
Mutate active site residues: catch it in the act



What to crystallize: add partner protein and/or substrate! (or anything else that will stabilize your protein)

Partner protein (if complex)
Peptide
DNA/RNA
ATP/GTP (or non-hydrolysable analogue)
Inhibitor
...
But how do you know what to add??

Study your protein!



... or cross your fingers and hope for the best

X family

Pol β **GCCGCGGGAAA**

(Pelletier '94) **CGGCGCCC**

Pol β **CGACTACGCGACAGCC**

(Batra '06) **GCTGATGCGC GTCGG**

A family

Bacillus **CGTACTACGAGAGA**

(Kiefer '98) **GCATGATGC**

T7 phage **GCTTTGCTGCCGG TCACGGTTCCCC**

(Doublie '98) **CGAAAACGACG GCCAGTGCCAAG**

Taq **CTGGTGCCGCGGGAAAA**

(Li '98) **GACCACGGCGCCC**

B family

RB69 **GCGCCTGACGAATGGACA**

(Franklin '01) **GCGGACTGCTTACT**

Y family

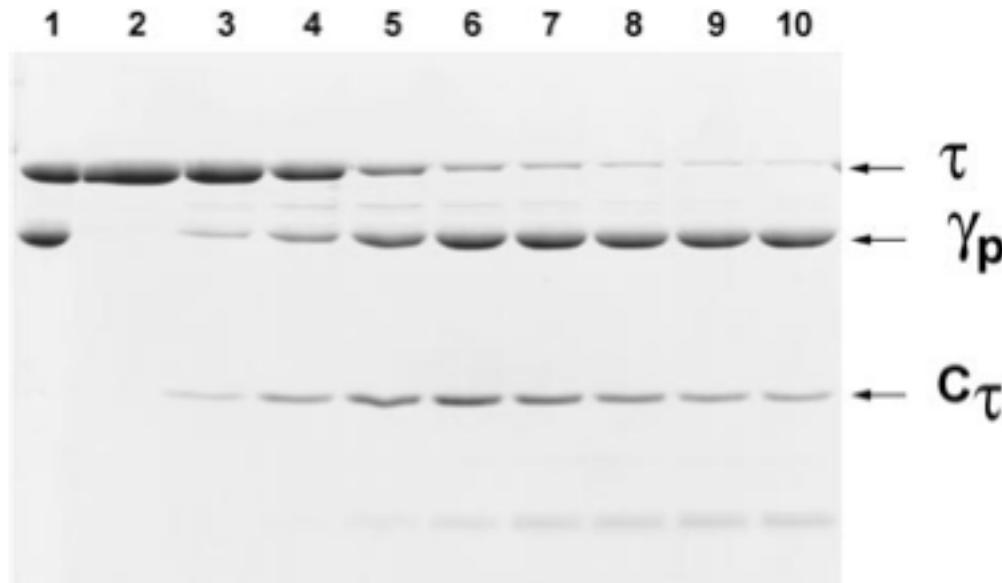
Dpo4 **CCCCCTTCCTGATTACTT**

(Ling '01) **GGGGGAAGGACTAA**

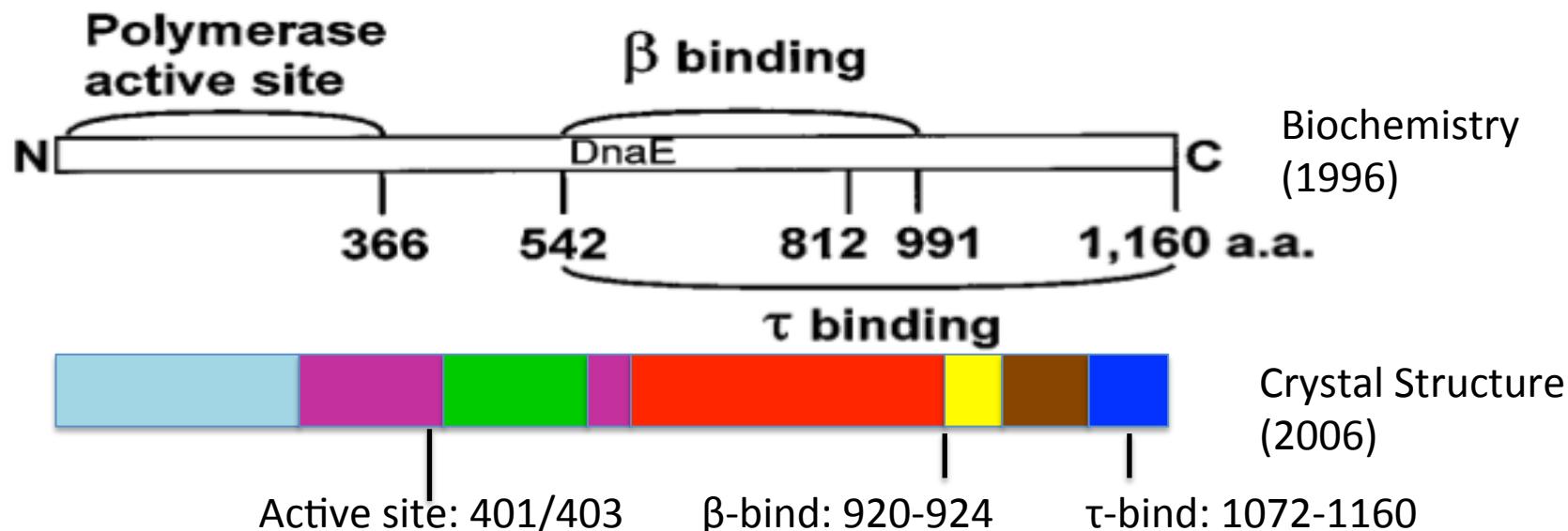
A = added during crystallization

A = not visible in structure

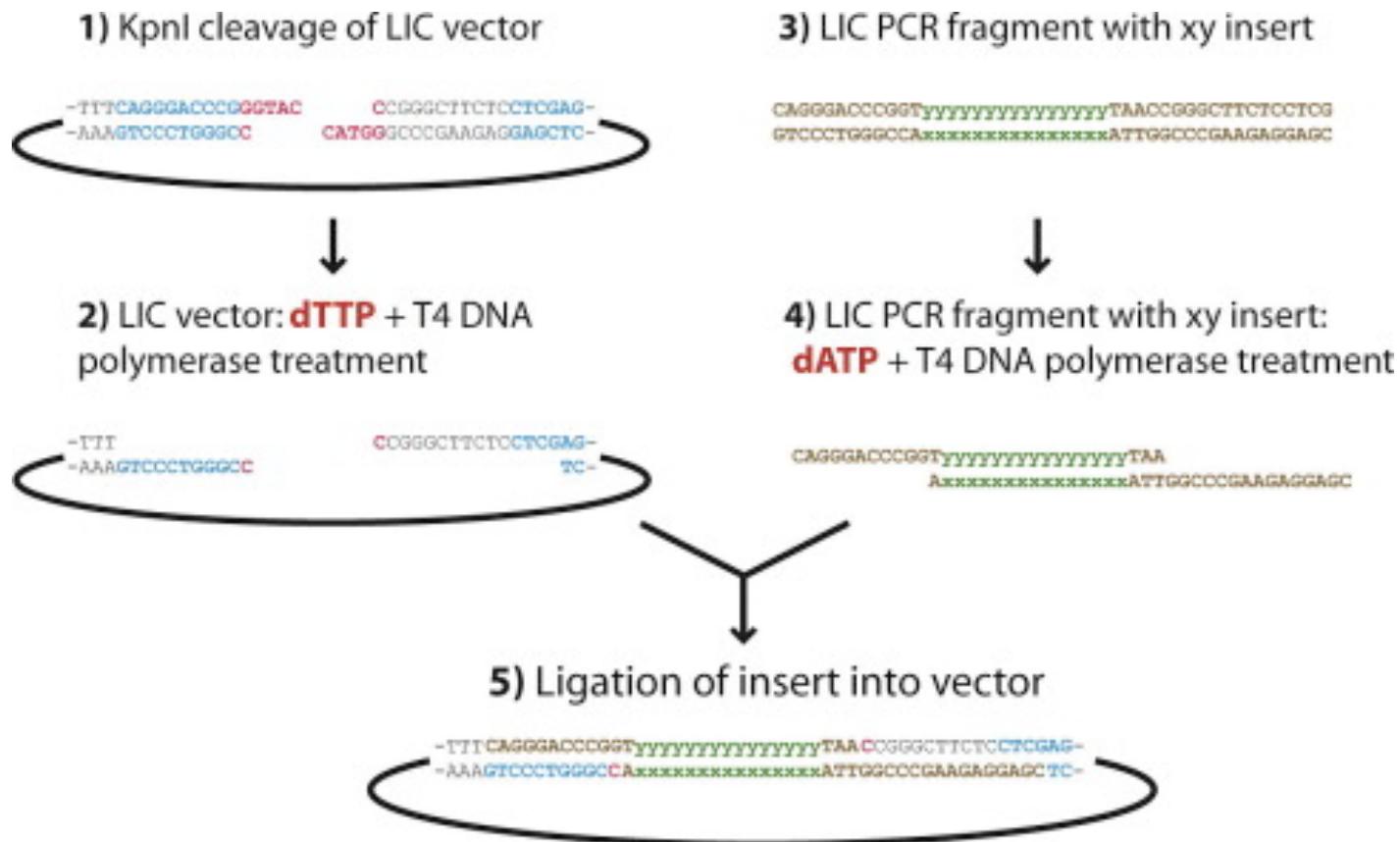
What to crystallize: DO NOT...



.. believe everything
that is published



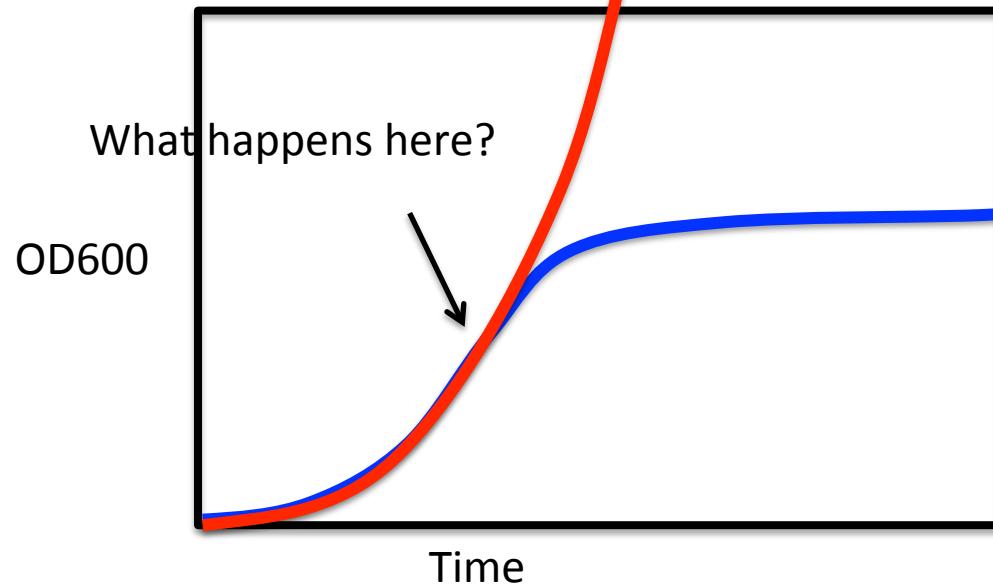
What to crystallize: how to clone



http://www2.lmb.internal/wiki/index.php/Lamers_lab -> cloning

Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 1990 Oct 25;18(20):6069-74. PubMed PMID: 2235490

Protein expression in E. coli



Studier, F. W. Protein production by auto-induction in high density shaking cultures.
Protein Expr Purif **41**, 207–234 (2005).

Table 10

Effect of magnesium on saturation density in 2×YT and terrific broth (TRB)

Growth medium	Source	A_{600}	pH
ZYM-505	Local	12.0	7.05
2×YT	Local	5.7	8.37
2×YT + 2 mM MgSO ₄	Local	8.3	8.44
TRB	Gibco/BRL	3.6	7.73
TRB + 2 mM MgSO ₄	Gibco/BRL	18.6	8.21
TRB	Local	12.5	8.06
TRB + 2 mM MgSO ₄	Local	18.1	8.18

E. coli is easily satisfied...

Need enough food: 2xTY (or even better: terrific broth)

Need Mg: 1mM MgSO₄

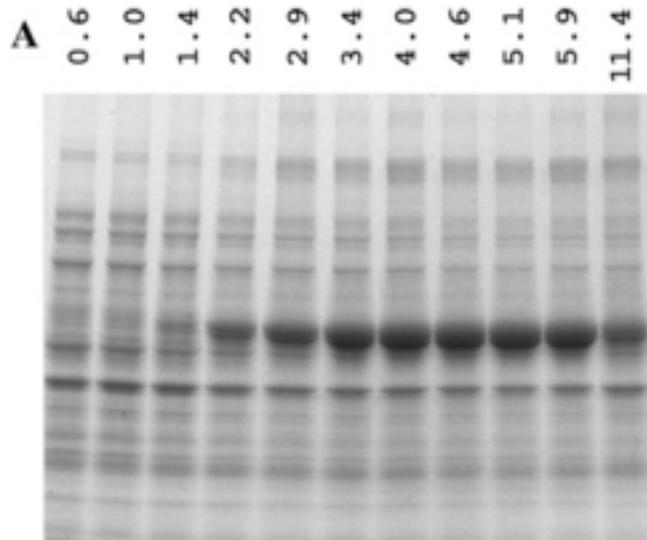
Need lots of O₂:use baffled flasks (Thomson UltraYield)

Prevent leaky expression: 1% glucose (and pLysS plasmid)

Don't like extra baggage (i.e your plasmids): do NOT do starter culture



Max protein production in 1-2 hrs...



Fool proof protocol:

Transform enough cell to plate out on 6 plates & grow overnight

Scrape ALL cells and inoculate 6 x 0.5 Ltr 2xTY + 1mM MgSO₄, 1% Glucose, Antibiotic

Grow 2-3 hrs at 37°C to OD₆₀₀=3-6 in BAFFLED UltraYield flasks (2.5 Ltr)

Add 1 volume of RT 2xTY (+Mg, Gluc, Antib, IPTG)

Express protein for 1-2 hrs @ 30°C

Harvest cells and freeze

=> 70-100g cells => 80-100mg protein (need ~10 mg to set up 2000 drops)

Sorry: I don't know much about protein expression in yeast, baculovirus or other systems

Don't ever say (or even think): "but this is how everyone does it"

Andrew Carter: Yeast to OD₆₀₀=80

Imre Berger: multi protein expression in baculovirus

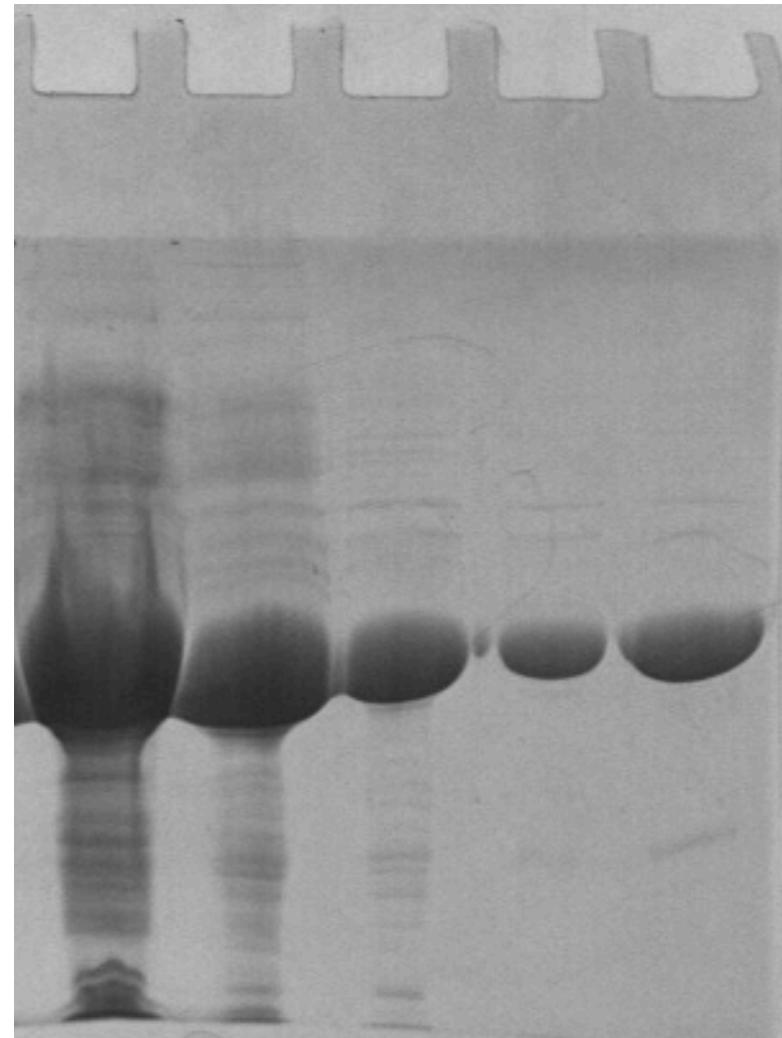
Wayne Hendrickson: SeMet expression for phasing

William Studier: using T7 phages for protein expression in E. coli

Protein purification

Tags
Hydrophobic
Ion-exchange
Affinity
Size exclusion
Buffer exchange
Protein storage

Lysate Phenyl SP Q S200



www.gelifesciences.com

> Service & support

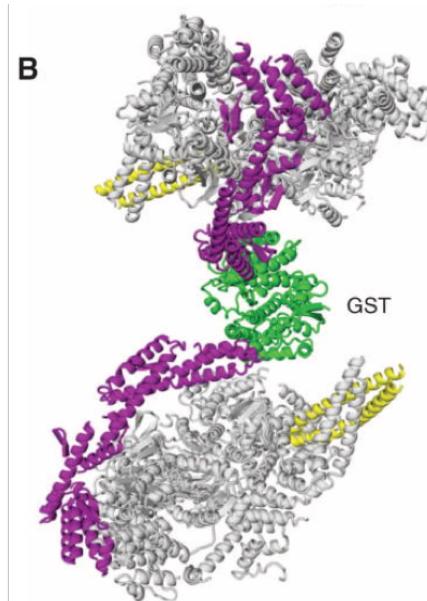
> Handbooks

Protein purification

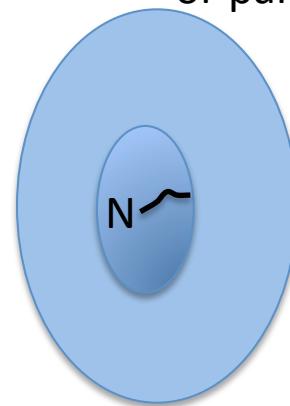
Tags

Hydrophobic
Ion-exchange
Affinity
Size exclusion
Buffer exchange
Protein storage

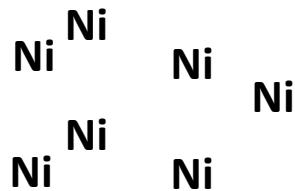
GST is a dimer!



N-terminus may be buried or part of active site!



Histrap columns can leak Ni into your protein



Protein purification

Tags

Hydrophobic

Ion-exchange

Affinity

Size exclusion

Buffer exchange

Protein storage

Do you really need to run a gel filtration run??

Only good if:

- You have lots of aggregates
- You have degradation products

⇒ Concentrating afterwards can do more harm than good



Protein purification

Tags

Hydrophobic

Ion-exchange

Affinity

Size exclusion

Buffer exchange

Protein storage

Instead of dialyzing try:

- * 5-10 fold dilution
- * Desalting column (30 minutes)
- * Changing pH by adding HCl or NaOH (within buffer capacity)

Protein purification

Tags

Hydrophobic

Ion-exchange

Affinity

Size exclusion

Buffer exchange

Protein storage

If you can, freeze your protein

Flash freeze protein in LN2. Use PCR tubes if needed

Thaw in hand, then store on ice

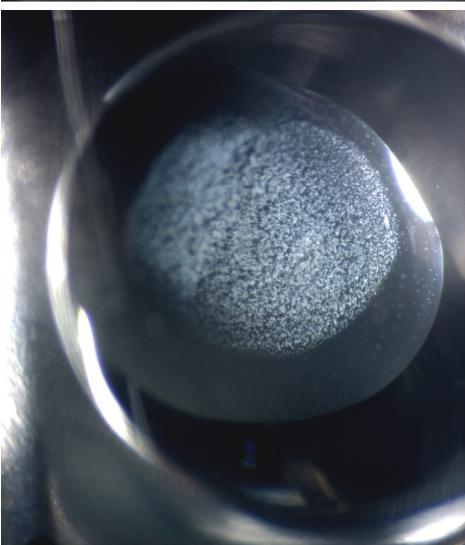
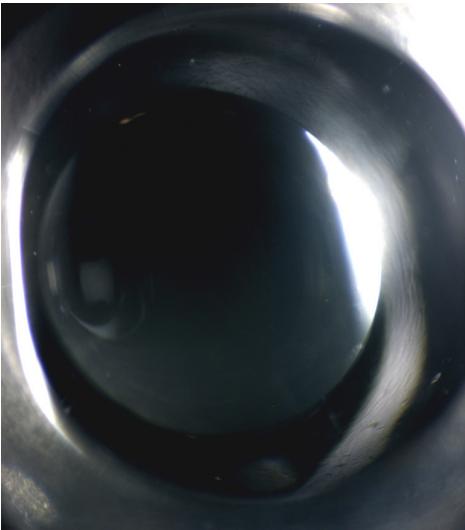
Find optimal storage buffer (use solubility screen)

Deng, J. *et al.* An improved protocol for rapid freezing of protein samples for long-term storage. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 203–204 (2004).

Protein characterization

1-10mg/ml
100nl/drop
Incubate 1-24 hrs

Solubility Screen

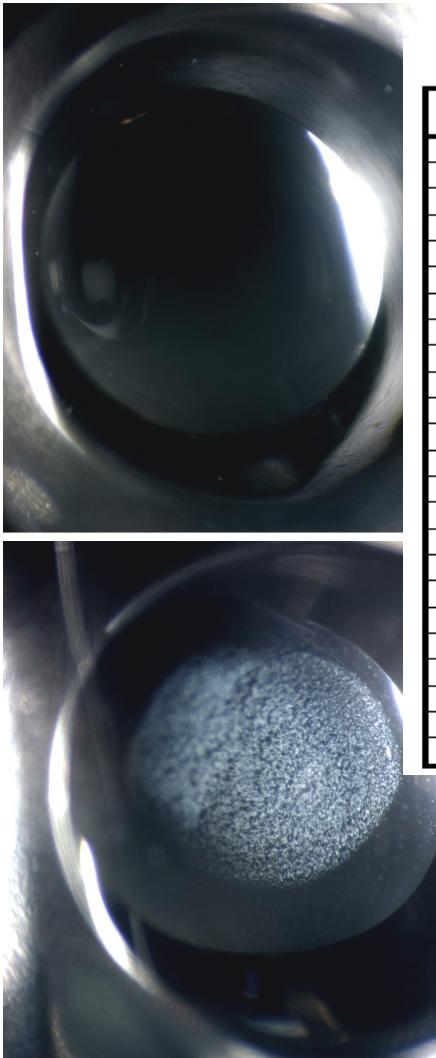


pH	Buffer (100mM)	Well			
		A1-B12	C1-D12	E1-F12	G1-H12
5.0	K Acetate	A1 20% TCA	C1	E1	G1
5.5	MES	A2	C2	E2	G2
6.0	ADA	A3	C3	E3	G3
6.5	MES	A4	C4	E4	G4
6.5	BisTris	A5	C5	E5	G5
6.5	KHPO ₄	A6	C6	E6	G6
7.0	PIPES	A7	C7	E7	G7
7.0	Imidazole	A8	C8	E8	G8
7.0	ADA	A9	C9	E9	G9
7.0	MOPS	A10	C10	E10	G10
7.5	HEPES	A11	C11	E11	G11
7.5	KHPO ₄	A12	C12	E12	G12
7.5	DIPSO	B1	D1	F1	H1
7.5	Tricine	B2	D2	F2	H2
8.0	HEPES	B3	D3	F3	H3
8.0	Bicine	B4	D4	F4	H4
8.0	Tris	B5	D5	F5	H5
8.5	Tris	B6	D6	F6	H6
8.5	TABS	B7	D7	F7	H7
8.5	TAPS	B8	D8	F8	H8
9.0	CAPSO	B9	D9	F9	H9
9.0	CHES	B10	D10	F10	H10
9.5	Glycine	B11	D11	F11	H11
10.0	CAPS	B12	D12	F12	H12 EMPTY

Protein characterization

1-10mg/ml
100nl/drop
Incubate 1-24 hrs

Solubility Screen



Protein 1 (4.5 mg/ml)					
pH	Buffer (100mM)	No Additive	NaCl (150mM)	Glycerol (5%V/V)	NaCl & Glyc
5.0	K Acetate	A1	C1	E1	G1
5.5	MES	A2	C2	E2	G2
6.0	ADA	A3	C3	E3	G3
6.5	MES	p	p	E4	p
6.5	BisTris	p	C5	E5	G5
6.5	KHPO ₄	p	C6	E6	G6
7.0	PIPES	p	C7	p	G7
7.0	Imidazole	p	C8	p	G8
7.0	ADA	A9	p	E9	G9
7.0	MOPS	p	C10	p	G10
7.5	HEPES	p	p	p	G11
7.5	KHPO ₄	p	C12	E12	p
7.5	DIPSO	p	D1	p	H1
7.5	Tricine	p	p	p	H2
8.0	HEPES	p	C13	p	H3
8.0	Bidme	p	C14	p	H4
8.0	Tris	p	p	p	H5
8.5	Tris	p	C15	p	H6
8.5	TABS	p	C17	p	p
8.5	TAPS	p	C18	p	H8
9.0	CAPSO	p	C19	p	H9
9.0	CHES	p	p	p	H10
9.5	Glydne	p	p	p	H11
10.0	CAPS	p	D12	p	H12

Protein 2 (10 mg/ml)			
No Additive	NaCl (150mM)	Glycerol (5%V/V)	NaCl & Glyc
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p
B1		p	H1
p	p	p	H2
p	p	p	H3
p	p	p	H4
B5		p	H5
B6	D6	F6	H6
B7	D7	F7	H7
B8	D8	F8	H8
B9		p	H9
B10		p	H10
B11		p	H11
B12		p	H12

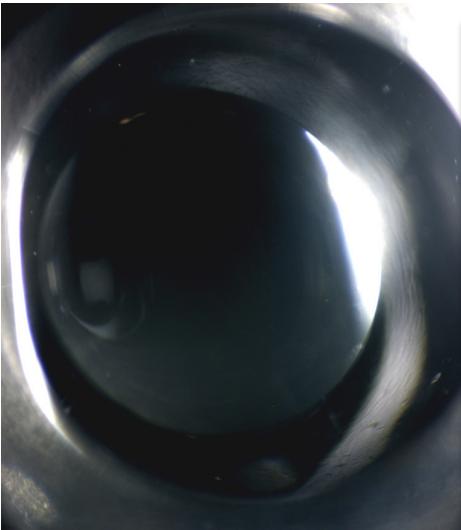
Protein 3 (15 mg/ml)			
No Additive	NaCl (150mM)	Glycerol (5%V/V)	NaCl & Glyc
p	p	p	p
p	p	p	p
p	C5		G3
p	C4		G4
A5	C5		G5
A6	C6	E6	G6
A7	C7	E7	G7
A8	C8		G8
A9	C9	E9	G9
p	C10		G10
p	C11		G11
A12	C12	E12	G12
B1	D1	F1	H1
p	p	p	H2
B3	D3	F3	H3
p	D4		H4
B5	D5	F5	H5
B6	D6	F6	H6
B7	D7	F7	H7
B8	D8	F8	H8
B9	D9	F9	H9
B10	D10	F10	H10
B11	D11		H11
B12	D12	F12	H12

See: http://www2.lmb.internal/wiki/index.php/Lamers_lab
 > Crystallography > Solubility screen
 Or Jena Bioscience JBScreen Solubility HTS

Protein characterization

1-10mg/ml
100nl/drop
Incubate 1-24 hrs

Solubility Screen



		Well			
		A1-B12	C1-D12	E1-F12	G1-H12
pH	Buffer (100mM)	No Additive	NaCl (150mM)	Glycerol (5%V/V)	NaCl & Glyc
9.0	CAPSO	B9	D9	F9	H9
9.0	CHES	B10	D10	F10	H10
9.5	Glycine	B11	D11	F11	H11
10.0	CAPS	B12	D12	F12	H12 EMPTY

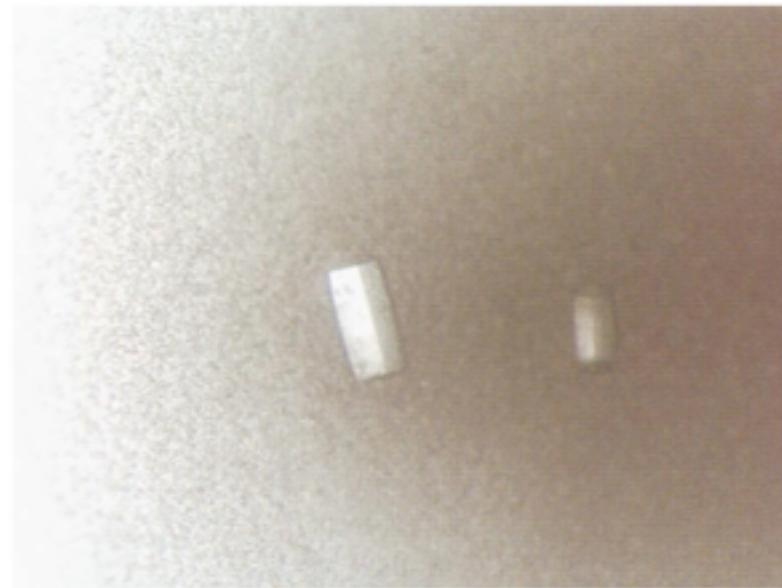
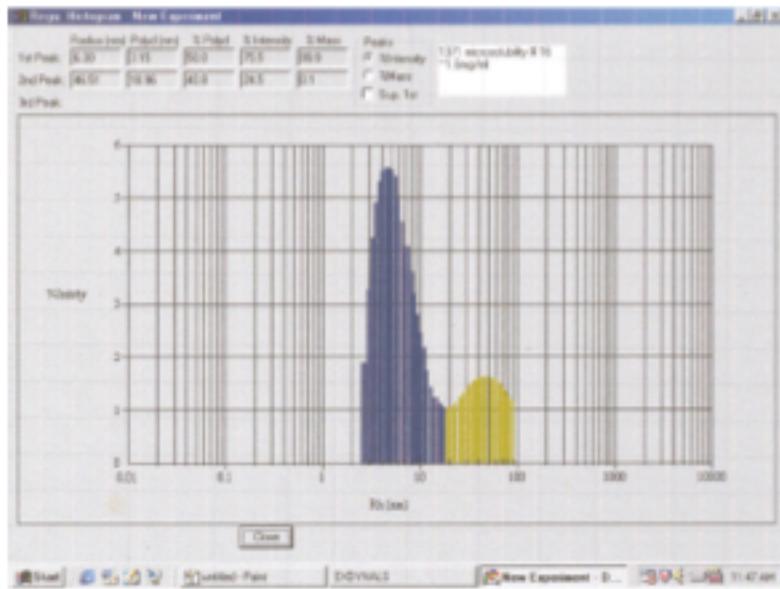
Now also available for:
activity assays
protein labeling
cross-linking
freezing

9.0	CAPSO	B9	D9	F9	H9
9.0	CHES	B10	D10	F10	H10
9.5	Glycine	B11	D11	F11	H11
10.0	CAPS	B12	D12	F12	H12 EMPTY

Protein characterization

Dynamic Light Scattering

5uL, 1-5mg/ml

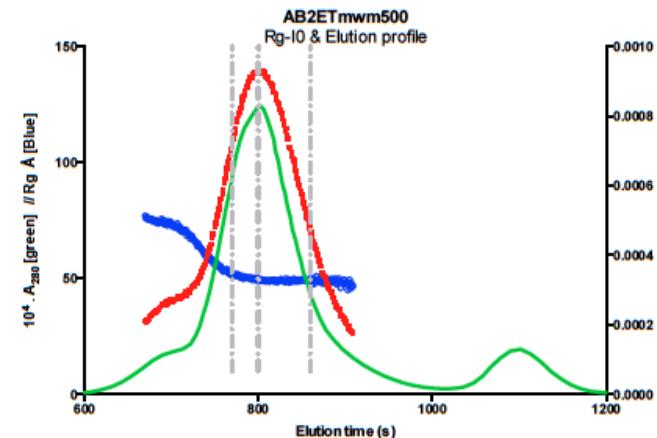
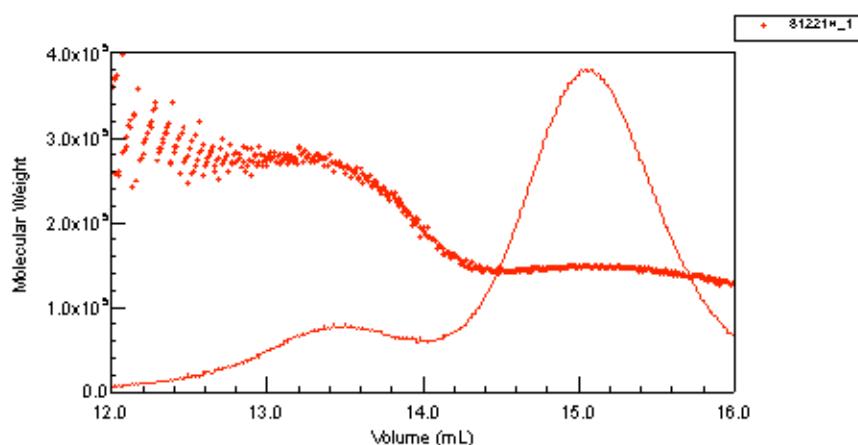
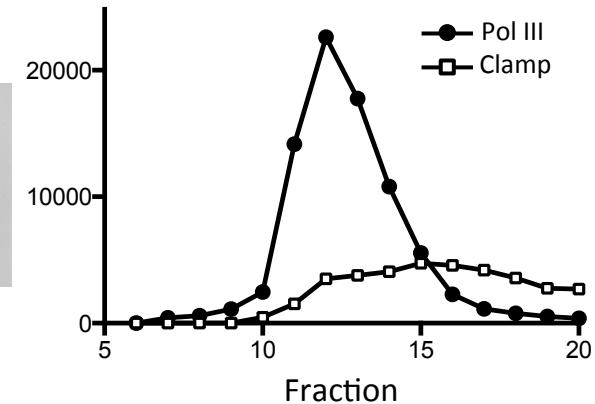
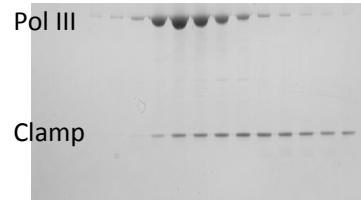
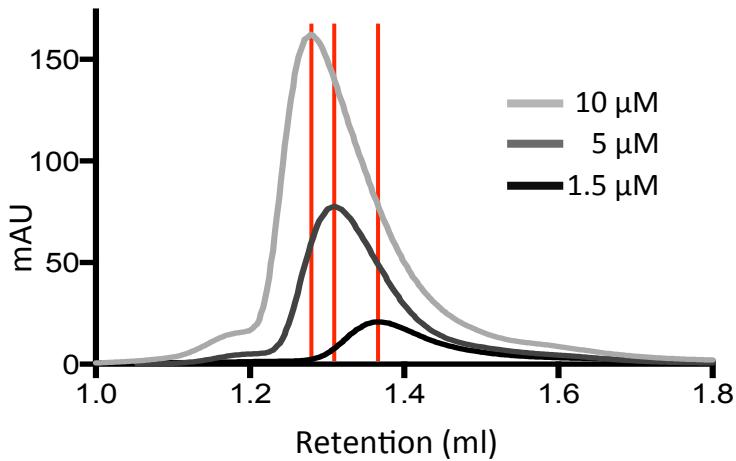


Monodisperse proteins give more crystal hits

Jancarik, J., Pufan, R., Hong, C., Kim, S.-H. & Kim, R. Optimum solubility (OS) screening: an efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. *Acta Cryst (2004). D60, 1670-1673 (2004)*

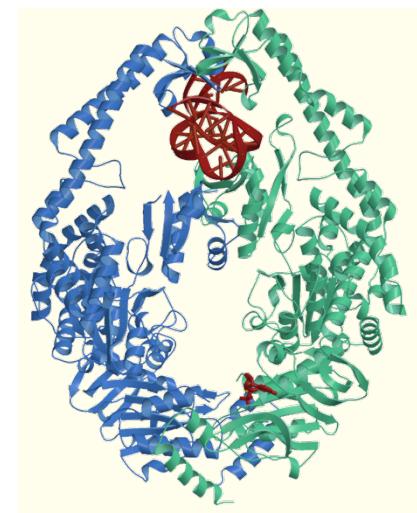
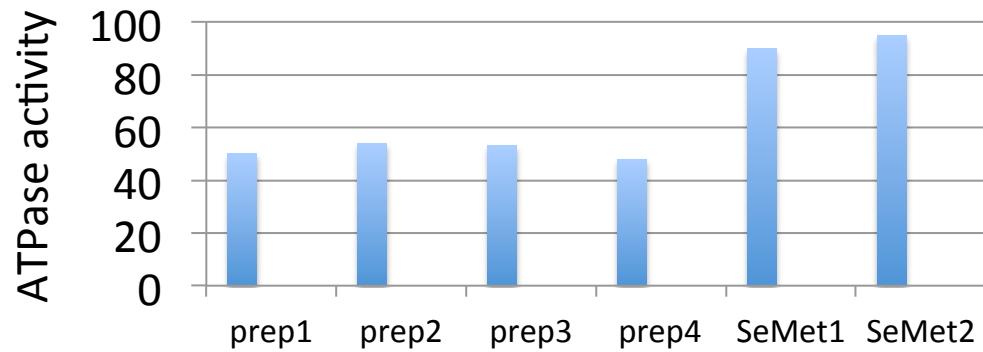
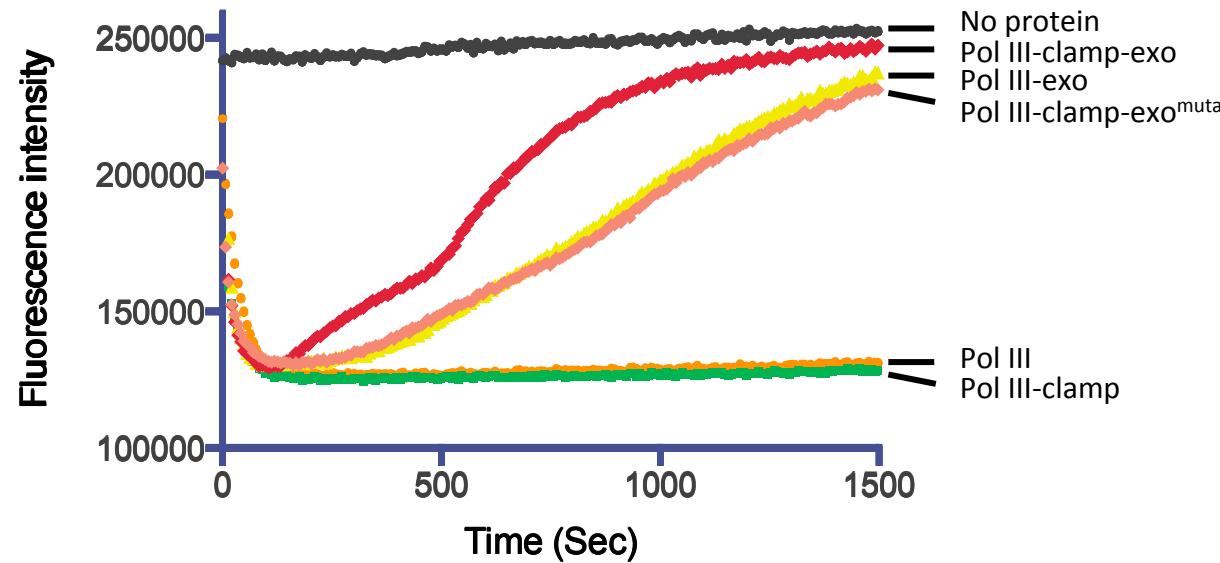
Protein characterization

Analytical gel filtration
SEC-MALS
SEC-SAXS



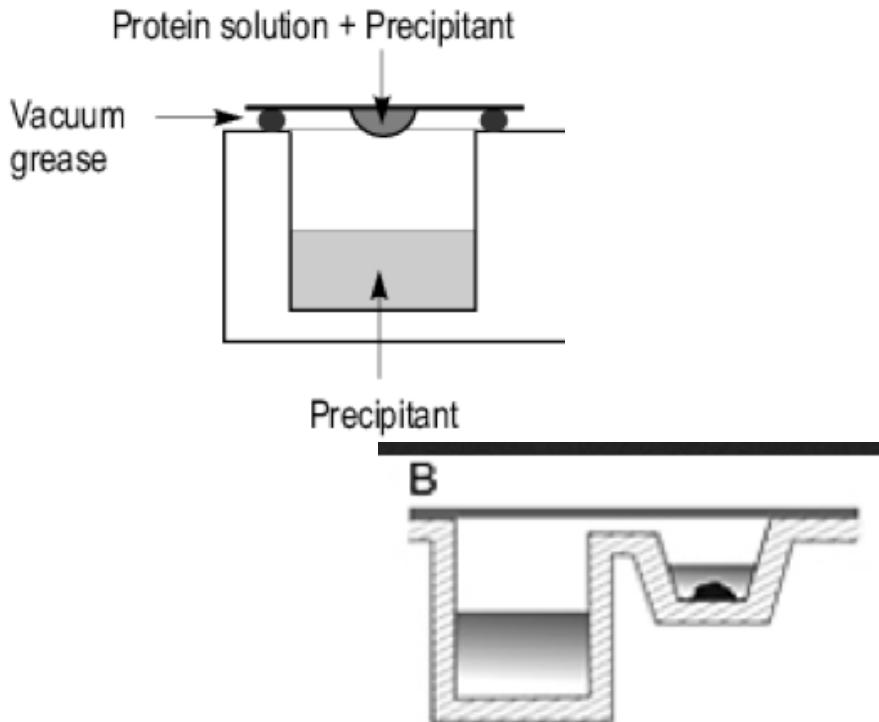
Protein characterization

Activity assay (Semet MutS)



Protein crystallization

Are robotic approaches the best way?



Diller DJ, Hol WG. An accurate numerical model for calculating the equilibration rate of a hanging-drop experiment. *Acta Crystallogr D Biol Crystallogr*. 1999 Mar;55(Pt 3):656-63. PubMed PMID: 10089462.

10 / 20 mg/ml

03/01/2012 15 16 17 18 19 20

1+1 µl	0.1		+	+		
(12)	0.2		+	+	+	+
	0.3	+	+	+	+	+
	0.4	+				
	15	16	17	18	19	20
1.5 + 0.5 µl / 0.1						
(11)	0.2				+	+
	0.3	+	+	+	+	+
	0.4	+				
	15	16	17	18	19	20
0.2 + 0.2 µl	0.1					
(6)	0.2	*		+	+++	++
	0.3	*	+++	+++	+++	+++
	0.4		+		+	
	15	16	17	18	19	20

Cryocrystallography (freezing)

Principles of cryocrystallography (1)

Freeze in liquid nitrogen or gas stream?

Beware of the layer of cold air over the liquid nitrogen (2)

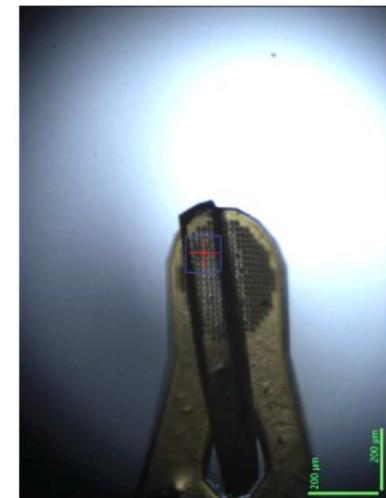
Test crystal at room temp: Mitegen crystal sleeve: MicroRT

Test different cryos & precipitant concentrations

Remove any mother liquor (3)

Dehydrate with PEGs

Dehydrate with controlled humidified air (4)



1. Garman, E. F. & Schneider, T. R. Macromolecular Cryocrystallography. *J. Appl. Cryst.* (1997). **30**, 211-237
2. Warkentin, M., Berejnov, V., Husseini, N. S. & Thorne, R. E. Hyperquenching for protein cryocrystallography. *J Appl Crystallogr* **39**, 805–811 (2006).
3. Pellegrini, E., Piano, D. & Bowler, M. W. Direct cryo cooling of naked crystals: are cryoprotection agents always necessary? *Acta Cryst* (2011). **D67**, 902-906
4. Russi, S. et al. Inducing phase changes in crystals of macromolecules: status and perspectives for controlled crystal dehydration. *J. Struct. Biol.* **175**, 236–243 (2011).



First DNA structure, 1953
(Watson & Crick)



First protein structure: Haemoglobin 1959
(Perutz & Kendrew)

In a difficult project you're not going to succeed because you do the same as everyone else, but because you do something **different**...

Recommended reading:

Ligation Independent Cloning

Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR).
Nucleic Acids Res. 1990 Oct 25;18(20):6069-74. PubMed PMID: 2235490

Protein expression

Studier, F. W. Protein production by auto-induction in high density shaking cultures.
Protein Expr Purif **41**, 207–234 (2005).

Protein engineering

Derewenda ZS. Application of protein engineering to enhance crystallizability and improve crystal properties. *Acta Crystallogr D Biol Crystallogr.*
2010 May;66(Pt 5):604-15. PubMed PMID: 20445236
SERP server <http://services.mbi.ucla.edu/SER/>

Protein purification

www.gelifesciences.com > Service & support > Handbooks

Hanging drop diffusion

Diller DJ, Hol WG. An accurate numerical model for calculating the equilibration rate of a hanging-drop experiment. *Acta Crystallogr D Biol Crystallogr.*
1999 Mar;55(Pt 3):656-63. PubMed PMID: 10089462

Cryo crystallography

Garman, E. F. & Schneider, T. R. Macromolecular Cryocrystallography. *J. Appl. Cryst* (1997). **30**, 211-237

Warkentin, M., Berejnov, V., Husseini, N. S. & Thorne, R. E. Hyperquenching for protein cryocrystallography. *J Appl Crystallogr* **39**, 805–811 (2006)

Pellegrini, E., Piano, D. & Bowler, M. W. Direct cryo cooling of naked crystals: are cryoprotection agents always necessary? *Acta Cryst* (2011). **D67**, 902-906

Russi, S et al. Inducing phase changes in crystals of macromolecules: status and perspectives for controlled crystal dehydration. *J. Struct. Biol.* **175**, 236-243 (2011)

When all goes well...

