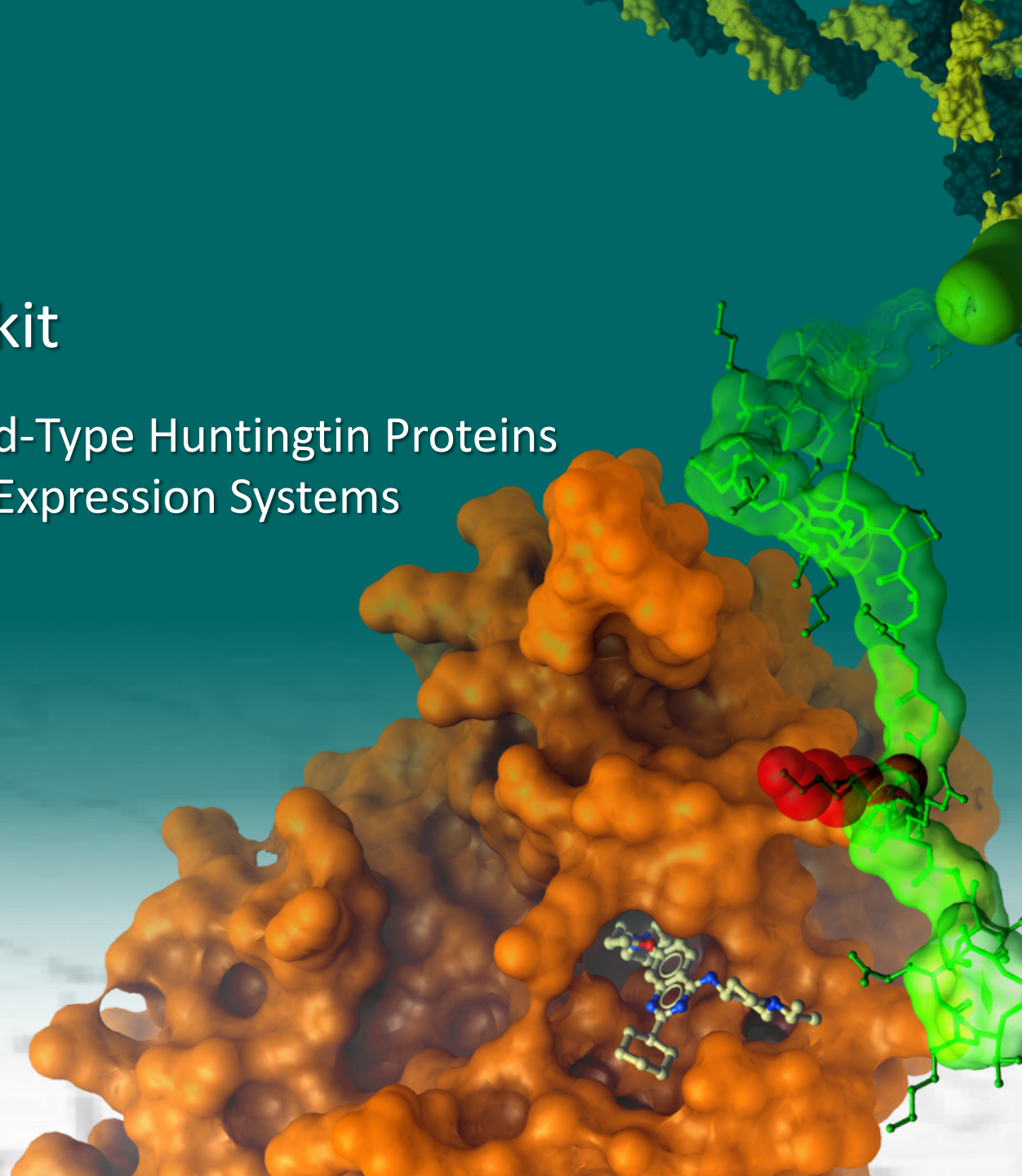
SGC

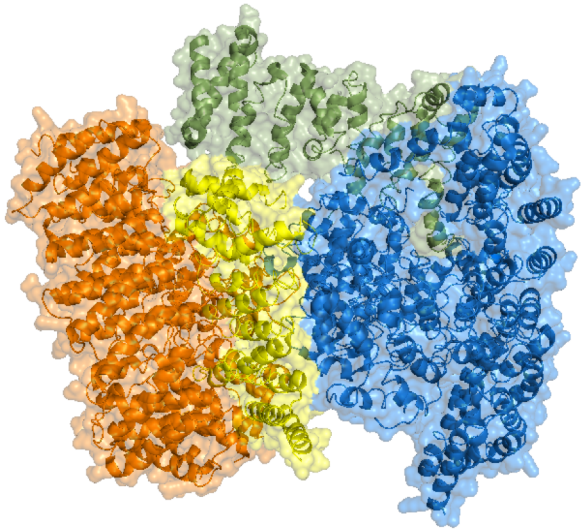
An Open Huntington's Disease Toolkit

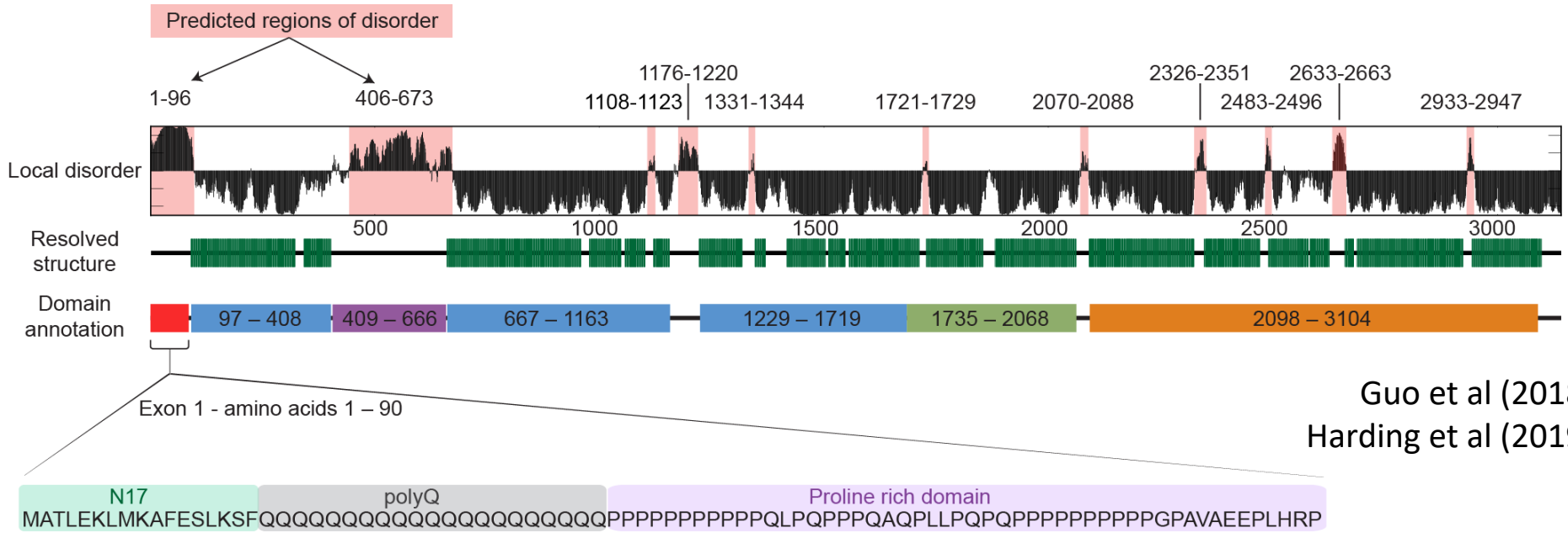
Design and Characterization of Mutant and Wild-Type Huntingtin Proteins Produced from a Toolkit of Scalable Eukaryotic Expression Systems

CAG Triplet Repeat Disorders - Gordon Research Conference
4th June 2019



CryoEM structure gives great insight into the HTT protein but there are still many questions





Guo et al (2018)
Harding et al (2019)

How does polyQ expansion of HTT affect protein structure? Can we determine a structure-function relationship?

What are the structures and functions of the intrinsically disordered regions of HTT?

Is HAP40 a constitutive binder of HTT? What is the function of HAP40?

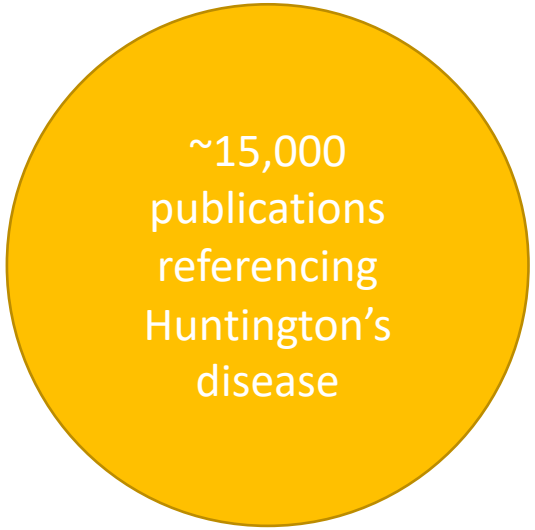
What other proteins does HTT bind to make physiologically relevant complexes?

A paucity of available biochemical tools have limited investigation of the HTT protein

1993: HTT gene identified

2006: Full-length HTT protein purification first published (Li *et al.)*

2018: No protein purification systems for large-scale full-length HTT production available through biorepositories.
Existing systems laid a great foundation but have issues with scalability, inflexibility of expression system, available polyQ lengths.

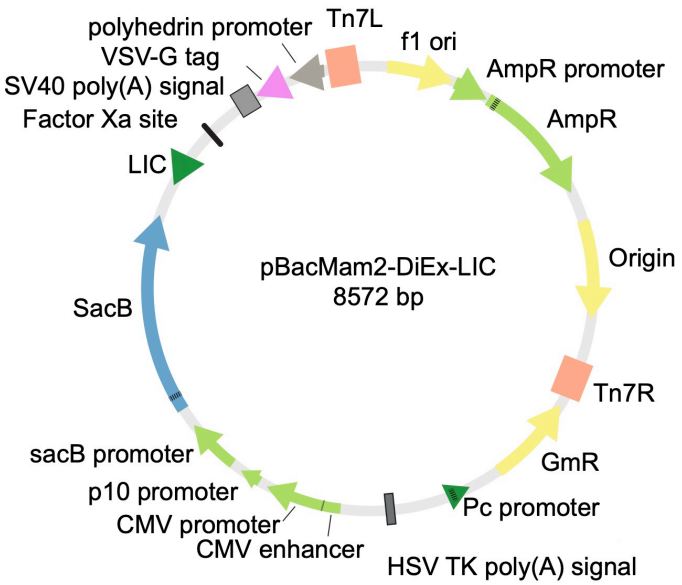


<50 publications referencing purified full-length HTT protein samples

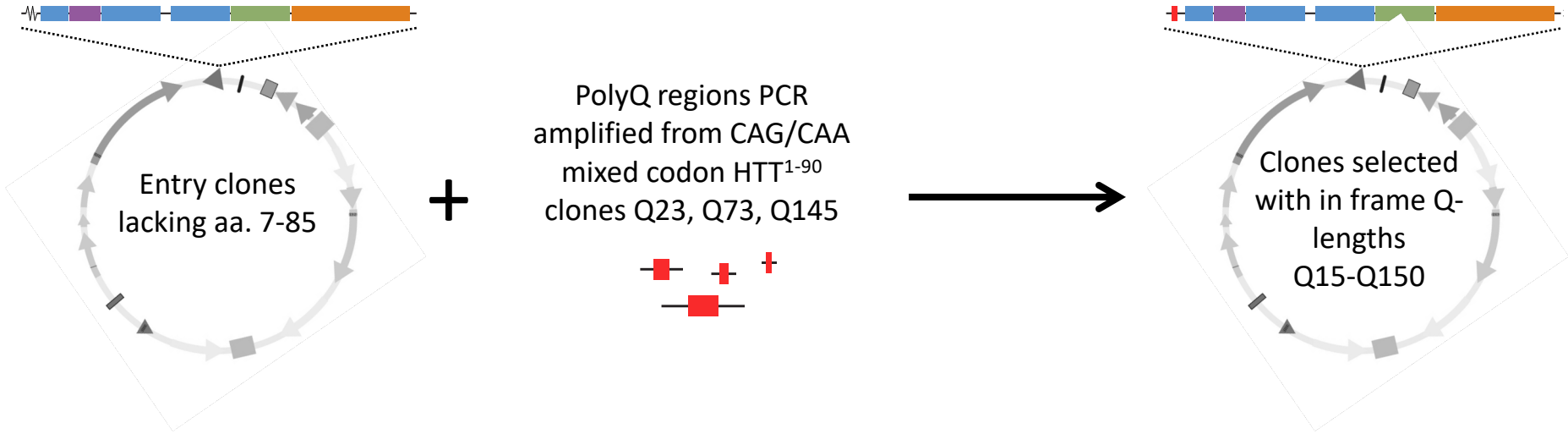
Data from PubMed search March 2019 for "Huntington's disease" & "purified huntingtin protein"

Aim: Design and characterise an open toolkit of Huntington's disease research resources for biochemical investigation of full-length HTT

pBacMam full-length HTT constructs cloned to freely share with HD community

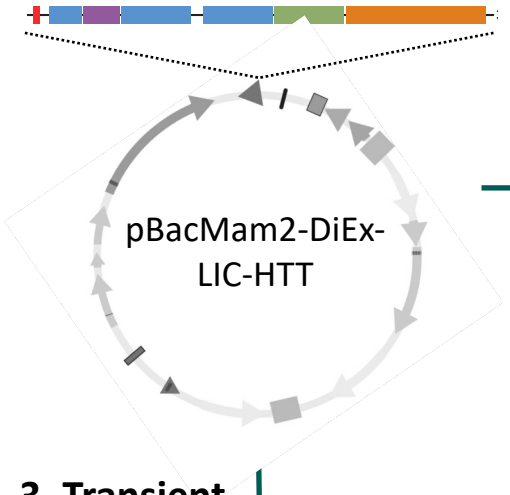


	N-terminal FLAG-tag	C-terminal FLAG-tag
General population	15, 19, 23	19, 23, 24, 25, 30
HD patients		36, 42, 48
Juvenile HD patients	51, 66, 73, 78, 79, 85	52, 54, 60, 66, 73, 79
Extreme expansions	139, 142, 145	109, 145



Peter Loppnau

Huntingtin expression in insect and human cells using pBacMam2-DiEx-LIC constructs



Bacmid DNA → Sf9 cell Transfection → Baculovirus recovery and amplification

1. Baculovirus expression in sf9 cells

2. Baculoviral transduction in EXPI293F

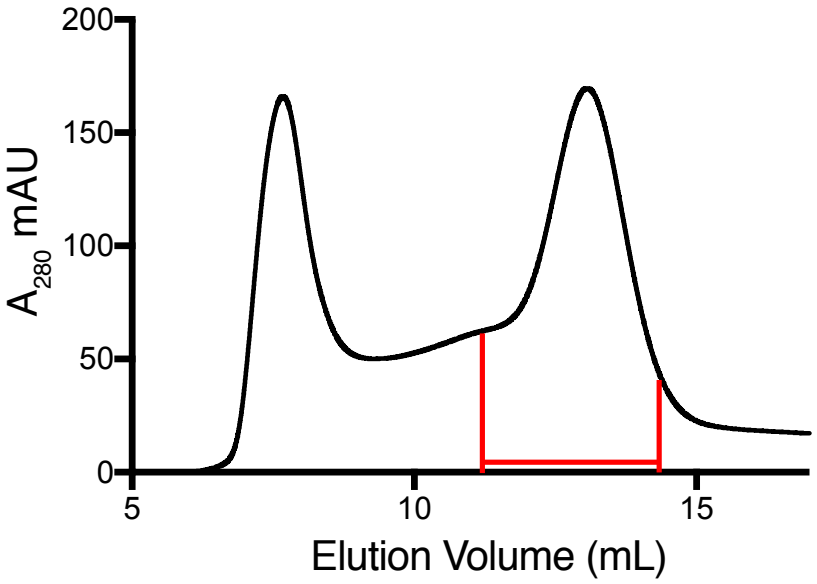
3. Transient transfection in EXPI293F

Suspension culture allows scalable production

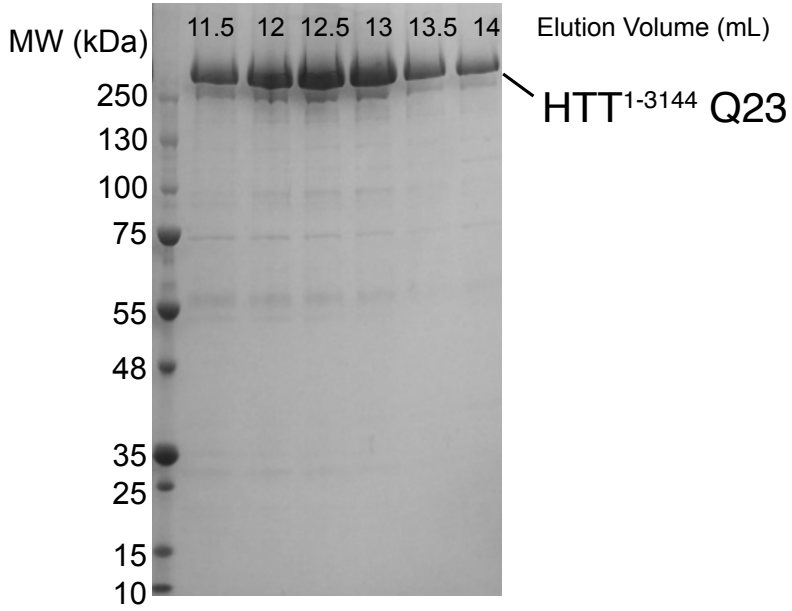
Alma Seitova and Ashley Hutchinson

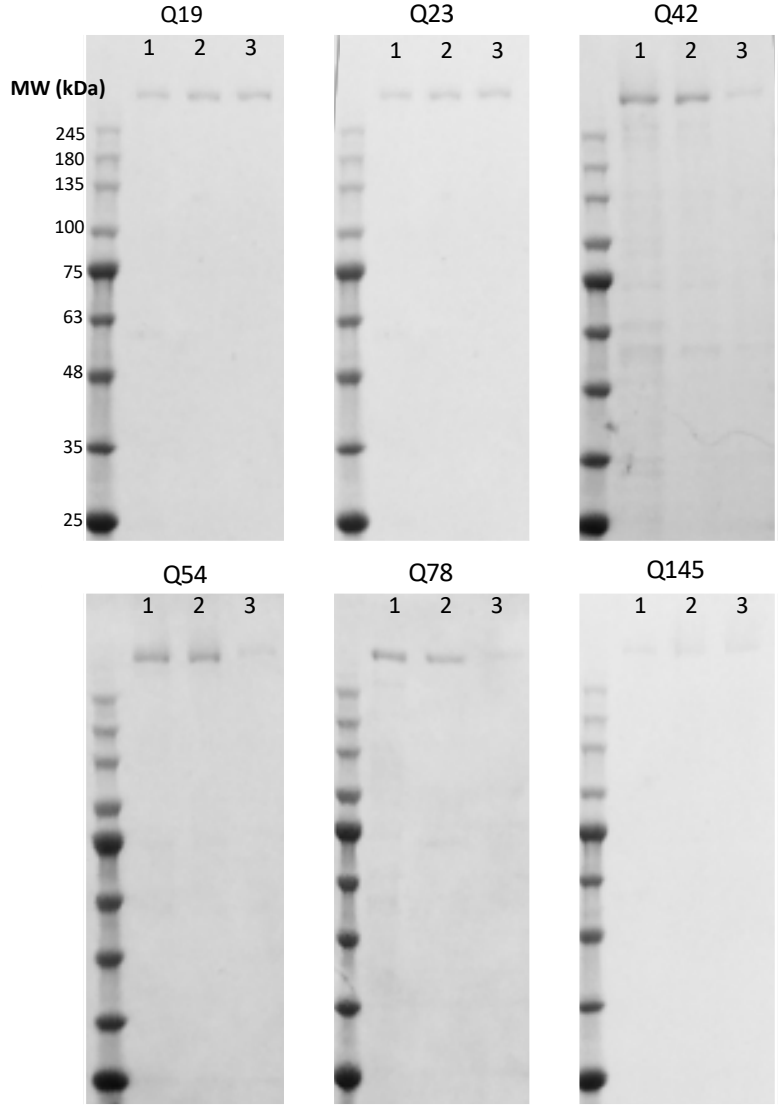
Full-length HTT of different Q-lengths can be trivially purified

Gel Filtration of insect HTT$^{1-3144}$ Q23



4-20% SDS-PAGE of monomer peak fractions





HTT proteins can be extracted to high purity, as determined by SDS-PAGE analysis, using FLAG affinity chromatography and gel filtration.

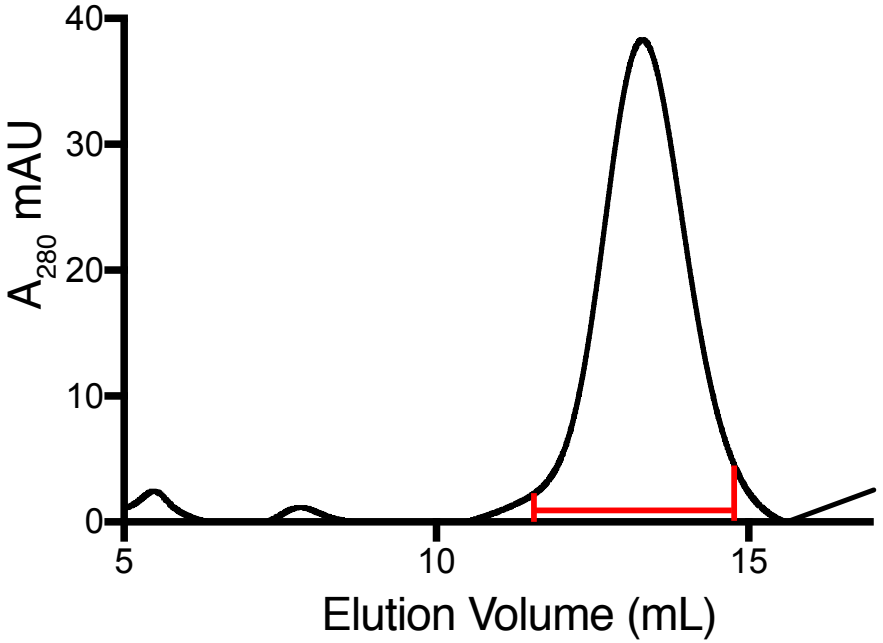
Yields for HTT$^{1-3144}$ Q23 ~1.6 mg/L Sf9 production.

The gel filtration profile of HTT indicates oligomeric heterogeneity.

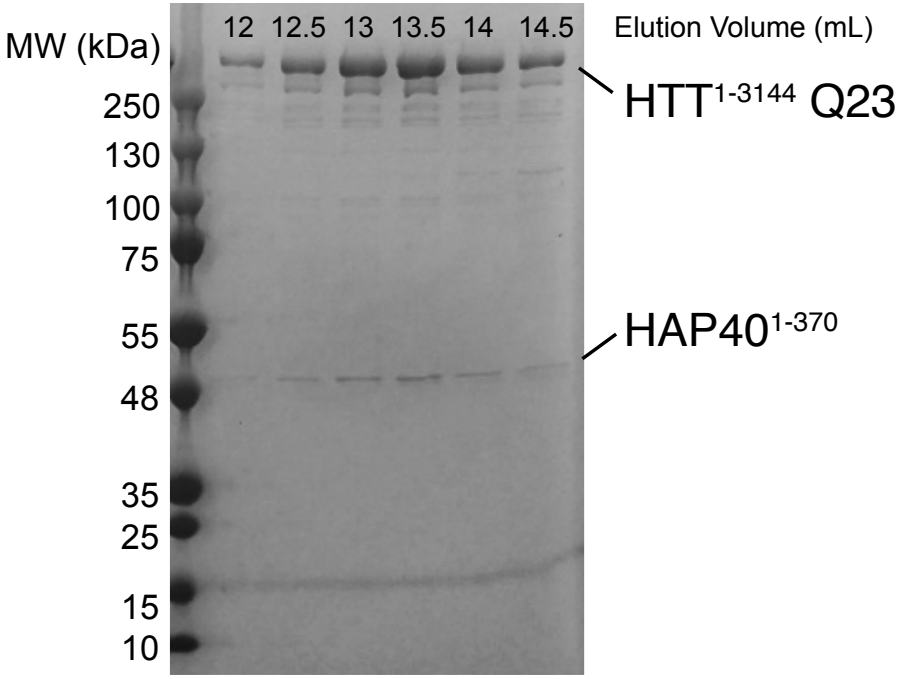
1 - baculovirus expression in sf9 insect cells
2 - transient transfection in mammalian EXPI293F cells
3 - transduction in mammalian EXPI293F cells

Purification of HTT-HAP40 complex from Sf9 baculovirus-mediated production

Gel Filtration of insect HTT$^{1-3144}$ Q23-HAP40$^{1-371}$



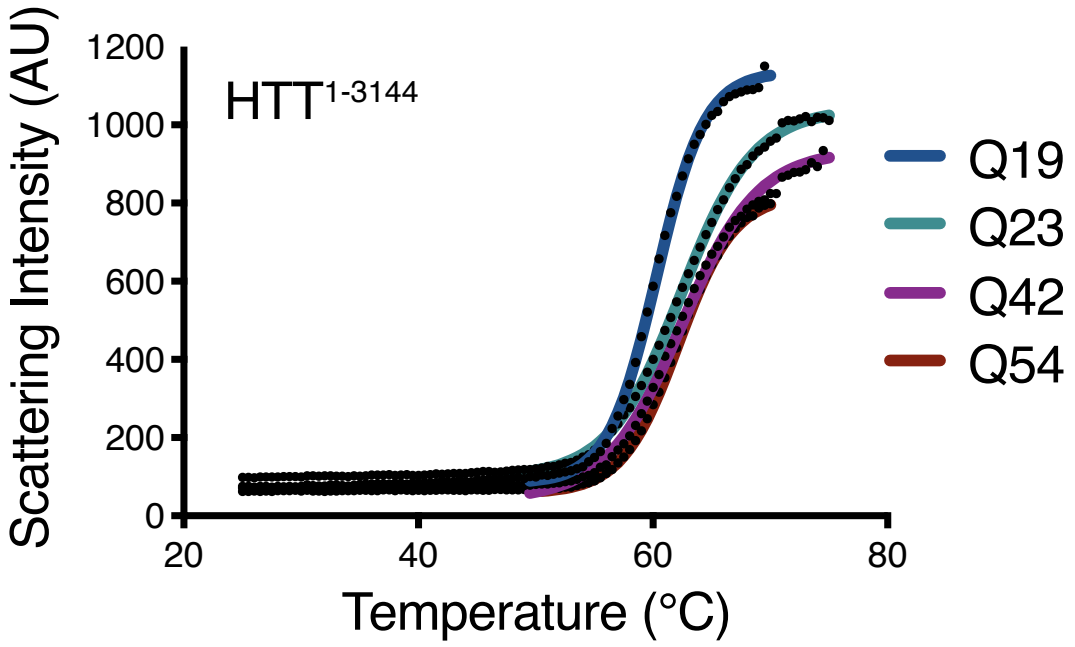
4-20% SDS-PAGE of complex peak fractions

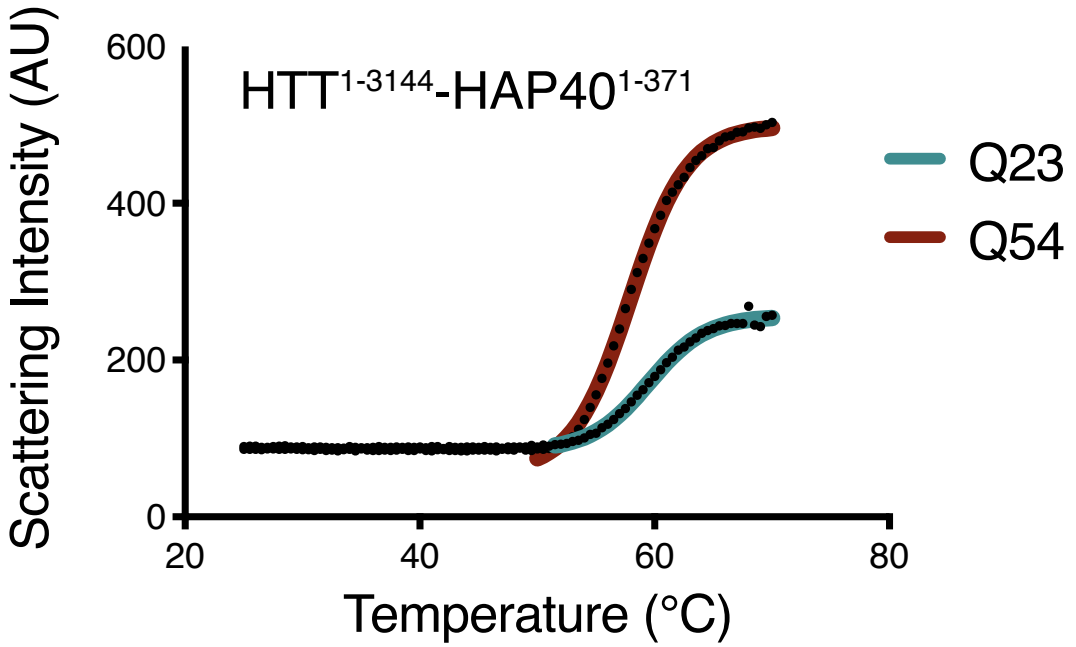


The constructs we have developed can be used to express and purify monodisperse HTT-HAP40 complex sample.

Formation of this complex by HTT produced in insect cells indicates that the protein expressed is correctly folded and functional with respect to formation of an important protein interaction

DSLS analysis shows HTT and HTT-HAP40 samples are stable and folded

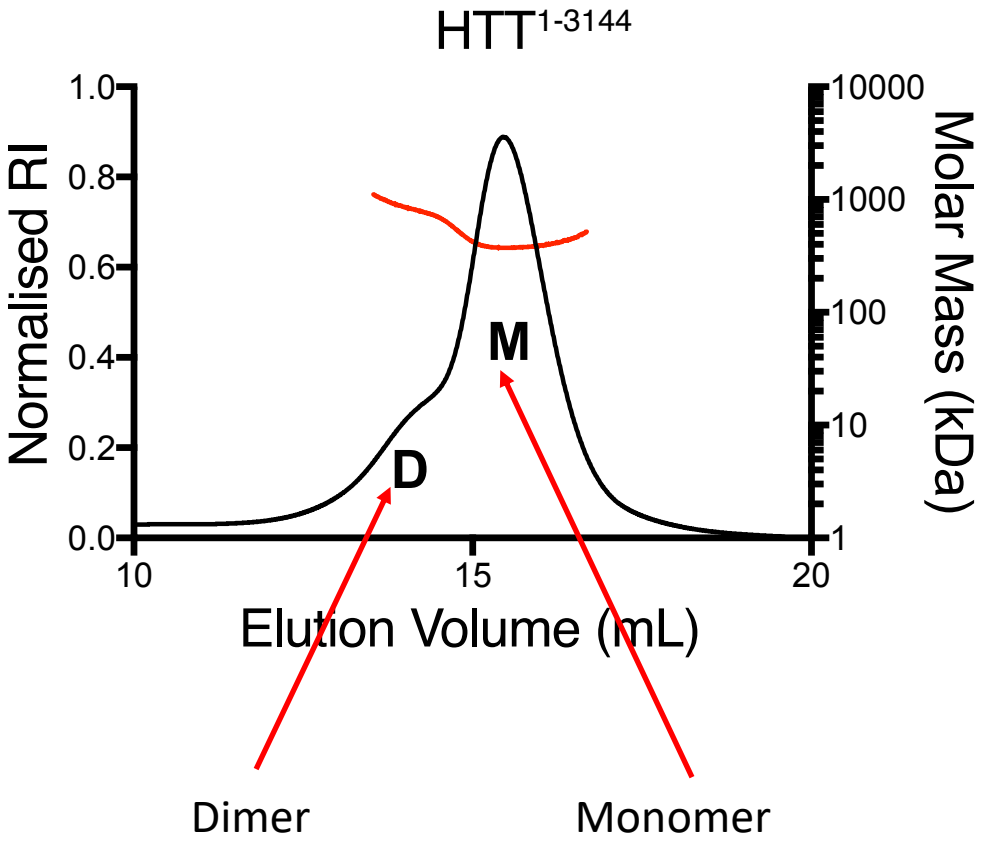


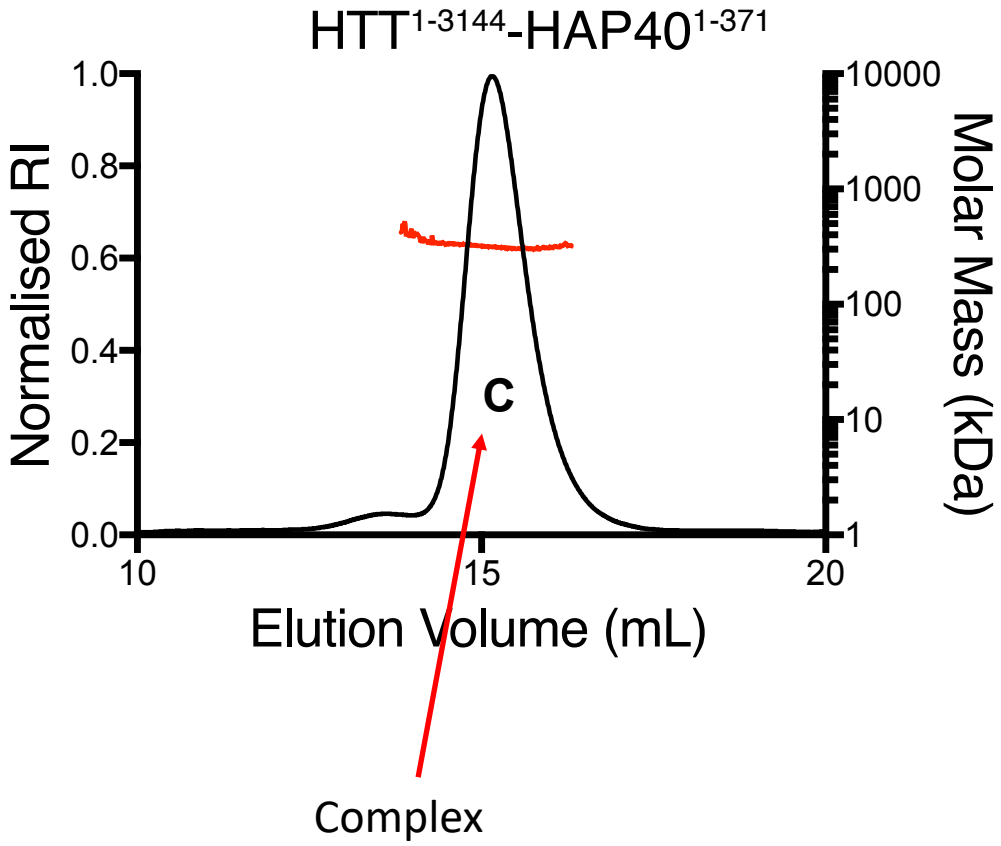


HTT samples are stable up to ~55 ˚C with sigmoidal thermal melting curves reflective of a folded globular protein.

In vitro thermal stability was independent of polyQ length.

MALS shows HTT self-associates but HTT-HAP40 complex is monodisperse





PTM motifs are similar for Sf9 compared to mammalian purified HTT

Purified HTT

Digest to peptides with trypsin, pepsin, chymotrypsin, lysarginase, mALP

ESI-MS

Peptide mapping

HTT$^{1-3144}$ Q23 from insect cells - Sf9

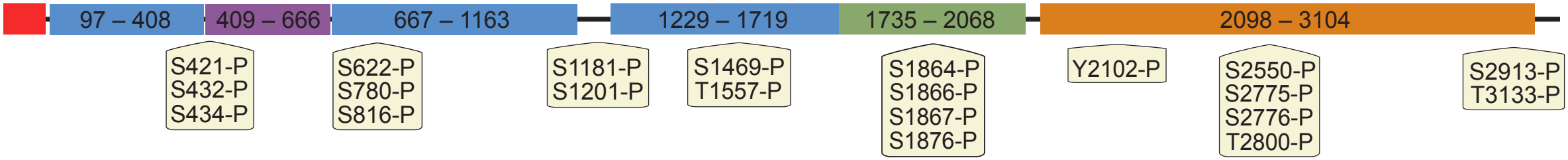
```
1    MATLEKLMKA FESLKSFQQQ QQQQQQQQQQ QQQQQQQQPP PPPPPPPPPQ LPQPPPQAQP LLPQPQPPPP PPPPPPGPAV AEEPLHRPKK ELSATKKDRV
101  NHCLTICENI VAQSVRNSPE FQKLLGIAME LFLLCSDDAE SDVRMVADEC LNKVIKALMD SNLPRLQLEL YKEIKKNGAP RSLRAALWRF AELAHLVRPQ
201  KCRPYLVNLL PCLTRTSKRP EESVQETLAA AVPKIMASFG NFANDNEIKV LLKAFIANLK SSSPTIRRTA AGSAVSICQH SRRTQYFYSW LLNVLLGLLV
301  PVEDEHSTLL ILGVLLTLRY LVPLLQQQVK DTSLKGSFGV TRKEMEVSPS AEQLVQVYEL TLHHTQHDHN NVVTGALELL QQLFRTPPPE LLQTLTAVGG
401  IGQLTAAKEE SGGRSRSGSI VELIAGGGSS CSPVLSRKQK GKVLLGEEEA LEDDSESRSD VSSSALTASV KDEISGELAA SSGVSTPGSA GHDIITEQPR
501  SQHTLQADSV DLASCDLTSS ATDGDEEDIL SHSSSQVSAV PSDPAMDLND GTQASSPISD SSQTTTEGPD SAVTPSDSSE IVLDGTDNQY LGLQIGQPQD
601  EDEEATGILP DEASEAFRNS SMALQQAHLL KNMSHCRQPS DSSVDKFVLR DEATEPGDQE NKPCRIKGDI GQSTDDDSAP LVHCVRLLSA SFLLTGGKNV
701  LVPDRDVRVS VKALALSCVG AAVALHPESF FSKLYKVPLD TTEYPEEQYV SDILNYIDHG DPQVRGATAI LCGTLICSIL SRSRFHVGDW MGTIRTLTGN
801  TFSLADCIPL LRKTLKDESS VTCKLACTAV RNCVMSLCSS SYSELGLQLI IDVLTLRNSS YWLVRTELLE TLAEIDFRLV SFLEAKAENL HRGAHHYTGL
901  LKLQERVLNN VVIHLLGDED PRVRHVAAAS LIRLVPKLFY KCDQGQADPV VAVARDQSSV YLKLLMHETQ PPSHFSVSTI TRIYRGYNLL PSITDVTMEN
1001 NLSRVIAAVS HELITSTTRA LTFGCCEALC LLSTAFPVCI WSLGWHCGVP PLSASDESRK SCTVGMATMI LTLSSAWFP LDLSAHQDAL ILAGNLLAAS
1101 APKSLRSSWA SEEEANPAAT KQEEVWPALG DRALVPMVEQ LFSHLLKVIN ICAHVLDDVA PGPAIKAALP SLTNPPSLSP IRRKGKEKEP GEQASVPLSP
1201 KKGSEASAAS RQSDTSGPVT TSKSSSLGSF YHLPSYLKLH DVLKATHANY KVTLDLQNST EKFGGFLRSA LDVLSQILEL ATLQDIGKCV EEILGYLKSC
1301 FSREPMMATV CVQQLLKTLF GTNLASQFDG LSSNPSKSQG RAQRLGSSSV RPGLYHYCFM APYTHFTQAL ADASLRNMVQ AEQENDTSGW FDVLQKVSTQ
1401 LKTNLTSVTK NRADKNAIHN HIRLFEPLVI KALKQYTTTT CVQLQKQVLD LLAQLVQLRV NYCLLDSDQV FIGFVLKQFE YIEVGQFRES EAIIPNIFFF
1501 LVLLSYERYH SKQIIGIPKI IQLCDGIMAS GRKAVTHAIP ALQPIVHDLF VLRGTNKADA GKELETQKEV VVSMLLRLIQ YHQVLEMFIL VLQQCHKENE
1601 DKWKRLSRQI ADIILPMLAK QQMHIDSHEA LGVLNTLFEI LAPSSLRPVD MLLRSMFVTP NTMASVSTVQ LWISGILAIL RVLISQSTED IVLSRIQELS
1701 FSPYLISCTV INRLRDGDST STLEEHSEGK QIKNLPEETF SRFLLQLVGI LLEDIVTKQL KVEMSEQQHT FYCQELGTLL MCLIHIFKSG MFRRITAAAT
1801 RLFRSDGCGG SFYTLDSLNL RARSMITTHP ALVLLWCQIL LLVNHTDYRW WAEVQQTPKR HSLSSTKLLS PQMSGEEEDS DLAAKLGMCN REIVRRGALI
1901 LFCDYVCQNL HDSEHLTWLI VNHIQDLISL SHEPPVQDFI SAVHRNSAAS GLFIQAIQSR CENLSTPTML KKTLQCLEGI HLSQSGAVLT LYVDRLLCTP
2001 FRVLARMVDI LACRRVEMLL AANLQSSMAQ LPMEELNRIQ EYLQSSGLAQ RHQRLYSLLD RFRLSTMQDS LSPSPPVSSH PLDGDGHVSL ETVSPDKDWY
2101 VHLVKSQCWT RSDSALLEGA ELVNRIPAED MNAFMMNSEF NLSLLAPCLS LGMSEISGGQ KSALFEAARE VTLARVSGTV QQLPAVHHVF QPELPAEPAA
2201 YWSKLNDLFG DAALYQSLPT LARALAQYLV VVSKLPSHLH LPPEKEKDIV KFVVATLEAL SWHLIHEQIP LSLDLQAGLD CCCLALQLPG LWSVVSSTEF
2301 VTHACSLIYC VHFILEAVAV QPGEQLLSPE RRTNTPKAIS EEEEEVDPNT QNPKYITAAC EMVAEMVESL QSVLALGHKR NSGVPAFLTP LLRNIIISLA
2401 RLPLVNSYTR VPPLVWKLGW SPKPGGDFGT AFPEIPVEFL QEKEVFKEFI YRINTLGWTS RTQFEETWAT LLGVLVTQPL VMEQEESPPE EDTERTQINV
2501 LAVQAITSLV LSAMTVPVAG NPAVSCLEQQ PRNKPLKALD TRFGRKLSII RGIVEQEIQA MVSKRENIAT HHLYQAWDPV PSLSPATTGA LISHEKLLLQ
2601 INPERELGSM SYKLGQVSIH SVWLGNSITP LREEEWDEEE EEEADAPAPS SPPTSPVNSR KHRAGVDIHS CSQFLLELYS RWILPSSSAR RTPAILISEV
2701 VRSLLVVSDL FTERNQFELM YVTLTELRRV HPSEDEILAQ YLVPATCKAA AVLGMDKAVA EPVSRLLEST LRSSHLPSRV GALHGVLYVL ECDLLDDTAK
2801 QLIPVISDYL LSNLKGIAHC VNIHSQQHVL VMCATAFYLI ENYPLDVGPE FSASIIQMCG VMLSGSEEST PSIIYHCALR GLERLLSEQ LSRLDAESLV
2901 KLSVDRVNVH SPHRAMAALG LMLTCMYTGK EKVSPGRTSD PNPAAPDSES VIVAMERVSV LFDRIRKGFP CEARVVARIL PQFLDDFFPP QDIMNKVIGE
3001 FLSNQQPYPQ FMATVVYKVF QTLHSTGQSS MVRDWVMLSL SNFTQRAPVA MATWSLSCFF VSASTSPWVA AILPHVISRM GKLEQVDVNL FCLVATDFYR
3101 HQIEEELDRR AFQSVLEVVA APGSPYHRLL TCLRNVHKVT TC
```

Sequence coverage is represented in black font with grey background.

Exon 1 (1-90)

Intrinsically disordered region (IDR – aa. 400-660)

HTT expressed in Sf9 insect cells retains reported phosphorylation PTMs



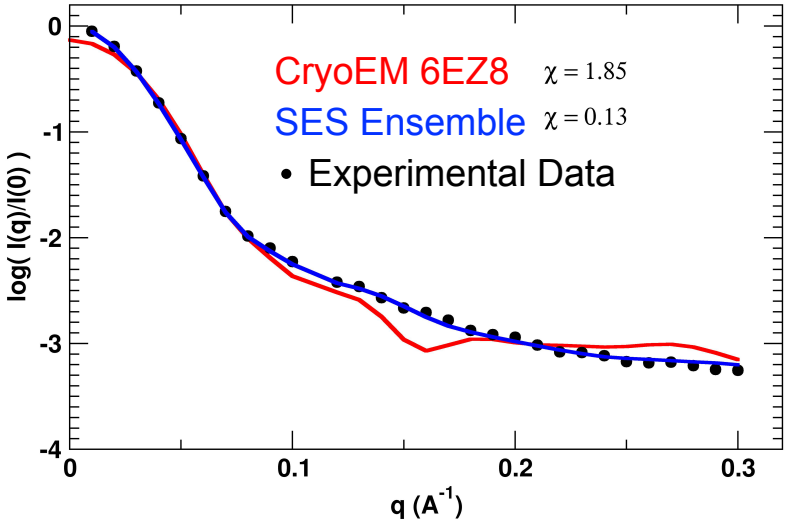
Suzanne Ackloo

Toolkit overview

- 28 constructs for N- and C- terminal FLAG-tagged full-length HTT expression
- 2 entry clones to permit generation of different polyQ expansion constructs
- Established protocols for purification of milligram quantities of HTT and HTT-HAP40 complexes from…
  - Baculovirus expression in sf9 insect cells
  - Transient transfection in mammalian EXPI293F cells
  - Transduction in mammalian EXPI293F cells
- Protein samples thoroughly characterised for…
  - Protein fold – DSLS
  - Monodispersity – analytical gel filtration and MALS
  - Function – HAP40 complex formation
  - Post-translational modification – motifs mapped by MS

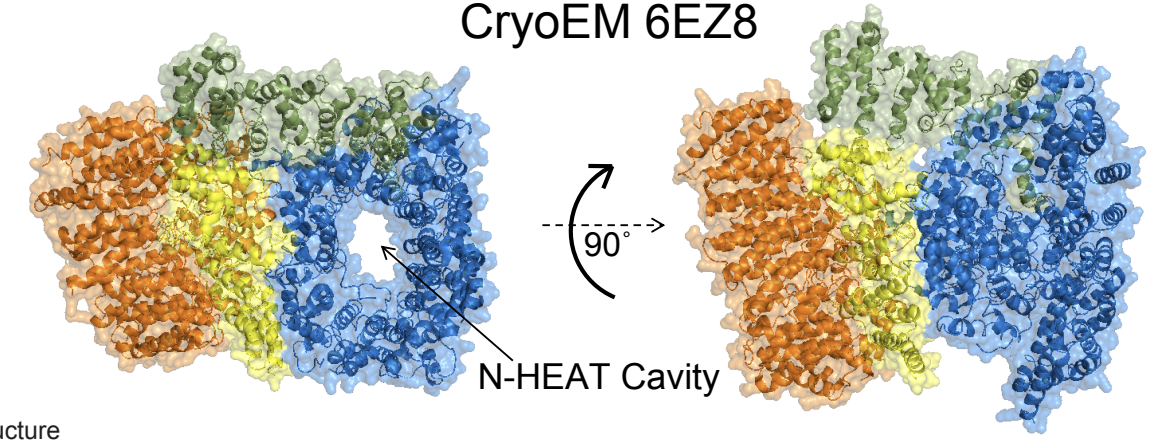
The toolkit allows analysis of the HTT protein with sample intensive methods – SAXS

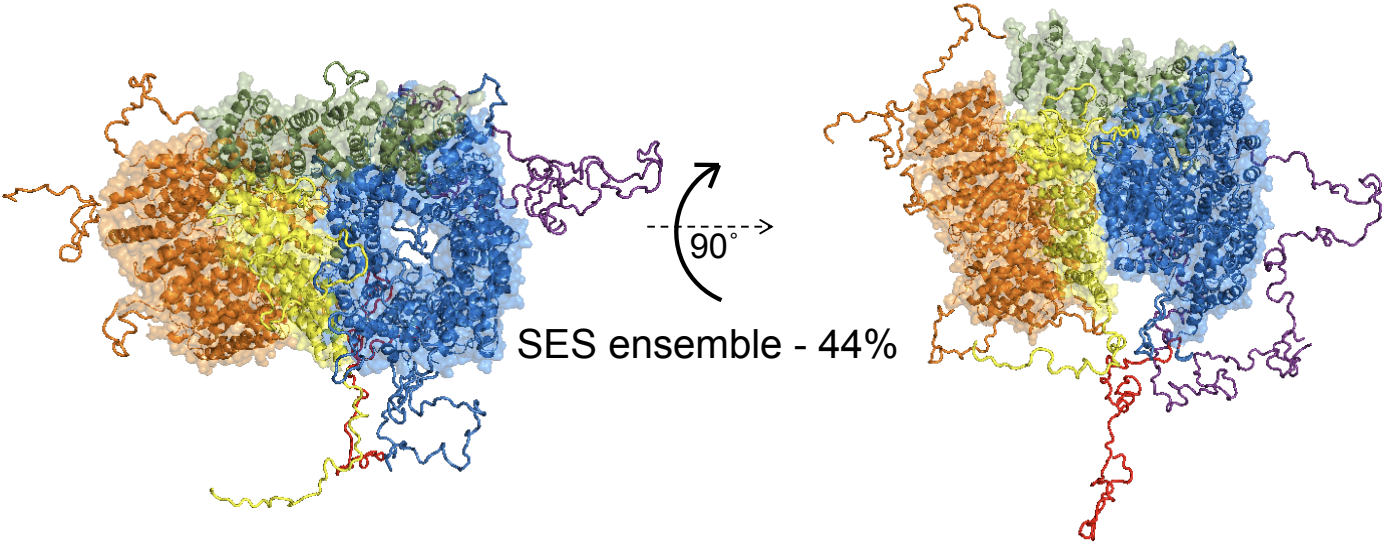
Experimental and theoretical SAXS profiles

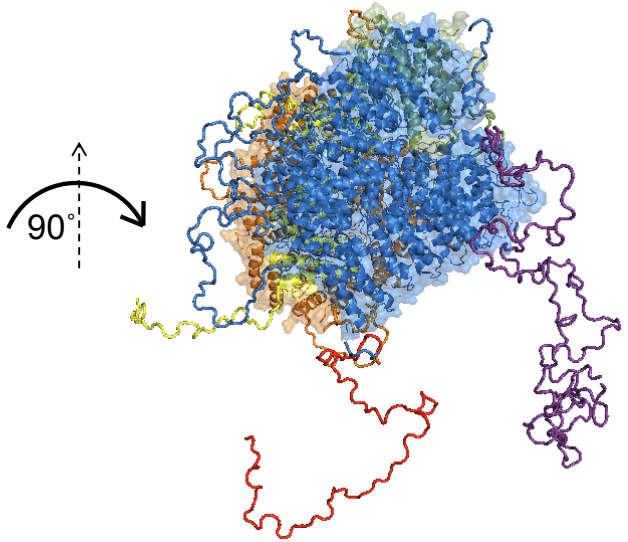


Protein sample	Rg (Å)	Rg (Å) (real)	Dmax (Å)	Mw (kDa)	Mw (kDa) (real)
HTT/HAP40	63.3+2.1	57.7+0.4	179	386	390

CryoEM 6EZ8







- Exon 1
- Intrinsically Disordered Region
- N-terminal HEAT
- Bridge
- C-terminal HEAT
- HAP40

pBacMam full-length HTT (and HAP40) constructs freely shared with HD community

SGC

Want to use these constructs in your own experiments?
All HTT pBMDEL plasmids and entry clones from this study are available through the Addgene repository

addgene
A better way to share plasmids

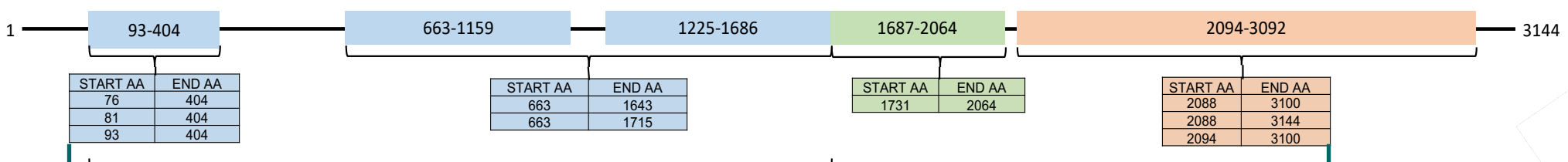
31	12	15
Total no. samples requested since December 2018	No. research groups requesting samples since December 2018	No. different constructs (Q-lengths) requested since December 2018

custom™ BIOLOGICS

Need HTT protein for your experiments?
Get in touch! We can help provide you with protein of any Q-length either via collaboration or CRO partnership

12	2
No. research groups HTT protein samples shared through collaboration since October 2017	CRO requests for purified HTT protein since January 2019

Adding to the toolkit: biophysically characterized HTT fragments



START AA	END AA
76	404
81	404
93	404

START AA	END AA
663	1643
663	1715

START AA	END AA
1731	2064

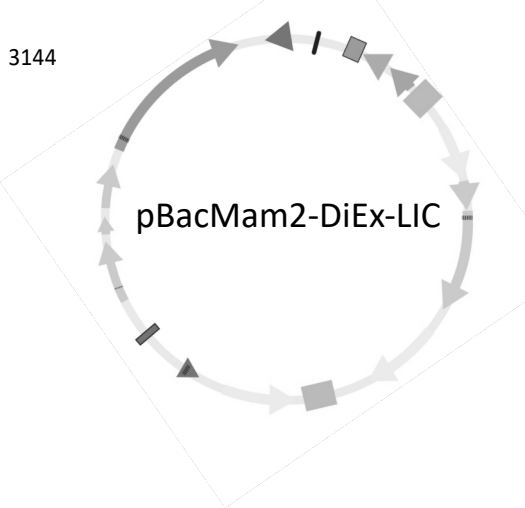
START AA	END AA
2088	3100
2088	3144
2094	3100

START AA	END AA
76	1643
76	1715
81	1643
81	1715

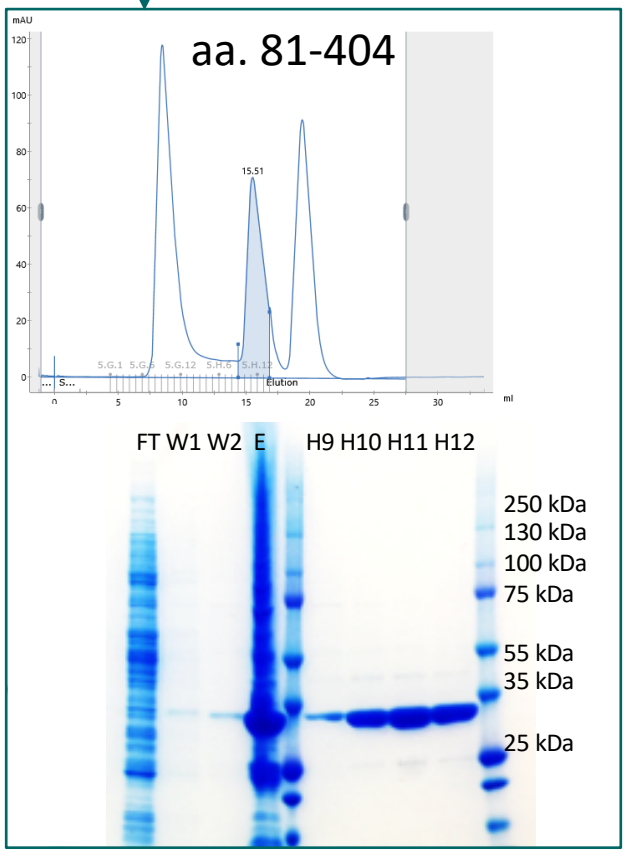
N-terminal HEAT

Bridge

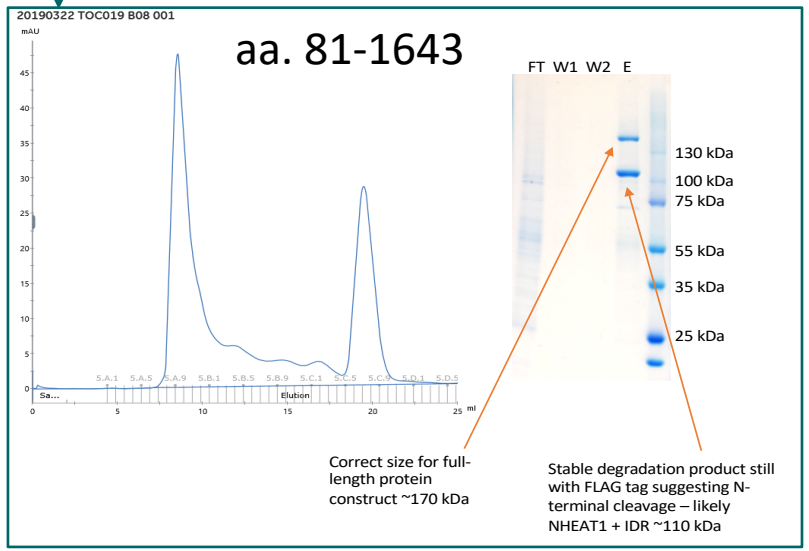
C-terminal HEAT



aa. 81-404



aa. 81-1643

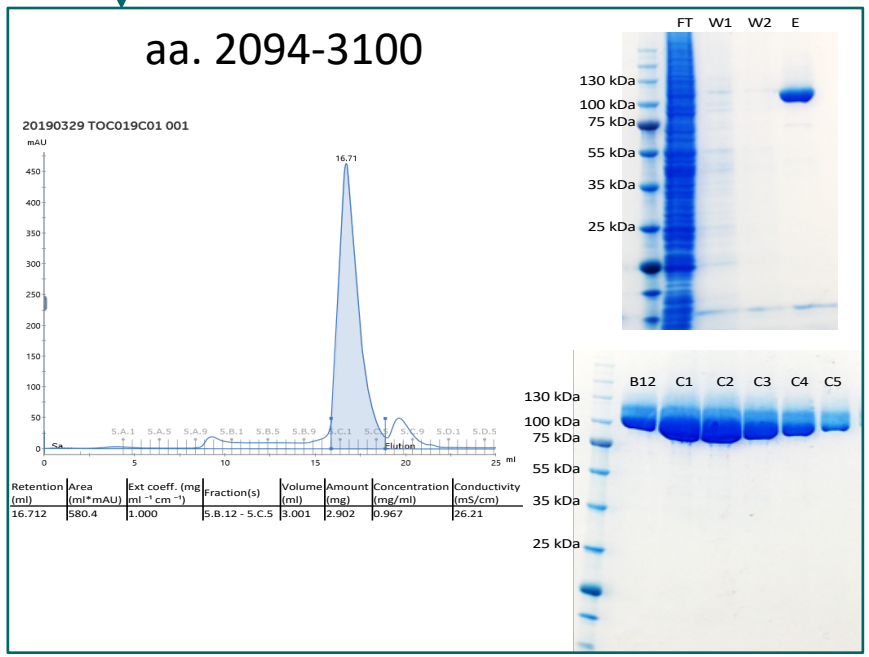


Crystallisation!

Tools for PPI mapping and other functional studies

Antibody generation

aa. 2094-3100



Publication of this work through my open lab notebook LabScribbles.com

1. Experiments completed in lab

2. Materials, methods, data and analysis uploaded to https://zenodo.org in open notebook community

3. Lay summary of experiment including discussion of context, aims and next steps plus links to Zenodo data upload

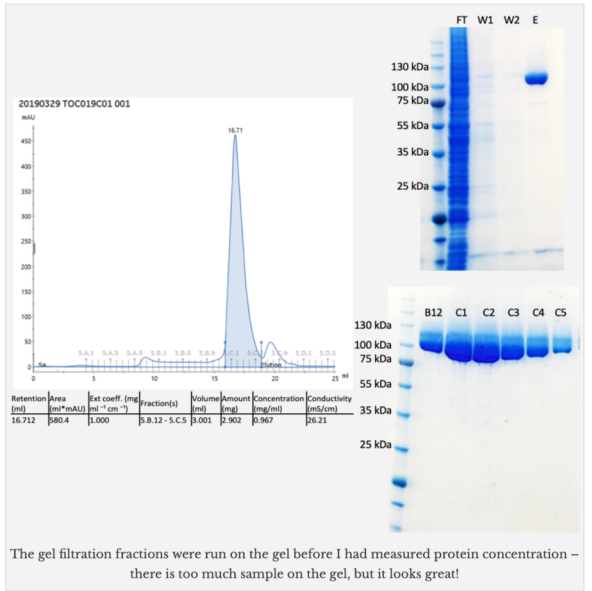
4. Share open notebook post via Twitter

Successful generation of fragments of the HTT protein and improving the purification procedure for the HTT-HAP40 complex

April 3, 2019 racheljaneharding Leave a comment Edit

It has now been almost 2 years since I set out to try and make fragments of the huntingtin protein which might be amenable to structural analysis with X-ray crystallography. X-ray crystallography is a fantastic method and allows us to see the molecules in very fine atomic detail which is important if we are to understand the intricacies of the elusive huntingtin protein molecule. It has been a long hard road with almost none of our extensive cloning efforts producing any expression constructs which made sufficient yields of protein. X-ray crystallography is a protein-expensive method so we need milligram (lots of protein). But finally, I have purified milligrams of different huntingtin fragment protein samples!!!

It should be made very clear that this has been a big team effort with cloner extraordinaire, Peter Loppnau, the eukaryotic production, Ashley Hutchinson and Alma Seitova as well as Linda Lin doing a lot of the heavy lifting on our cloning and eukaryotic production pipeline so I am very grateful for all of their hard work. Turns out that my construct design was fine but the expression vector we used made a huge difference (this is the piece of DNA which we insert different parts of the huntingtin gene). Switching from pFBOH-MHL to pBMDEL gave us great yields! Here is the C-terminal HEAT domain protein I purified – so much pure protein! All of the other data can be found on Zenodo.



The gel filtration fractions were run on the gel before I had measured protein concentration – there is too much sample on the gel, but it looks great!

Acknowledgements

SGC and UHN
Jacob McAuley
Claudia Alvarez
Peter Loppnau
Ashley Hutchinson
Alma Seitova
Mani Ravichandran
Levon Halabelian
Suzanne Ackloo
Shili Duan
Sasha Lemak
Linda Lin
Cheryl Arrowsmith
Aled Edwards

Ulm University
Stefan Kochanek

EPFL
Hilal Lashuel
Driss Boudeffa

CHDI Foundation
Leticia Toledo-Sherman
Matt Lee
Liz Doherty

Oxford University
Justin Deme
Susan Lea
Bass Hassan

McMaster University
Ray Truant
Tam Maiuri

Advanced Proton Source
Xiaobing Zuo
Lixin Fan

John Hopkins University
Christopher Ross
Tamara Ratovitski

Western Washington University
Jeff Carroll

A*STAR
Mahmoud Pouladi

Washington University St Louis
Alex Holehouse

Sick Kids Hospital Toronto
Chris Pearson
Terence Gall-Duncan
Babak Koucheki

Max-Delbrück-Center for Molecular Medicine
Erich Wanker
Philipp Trepte
Anne Ast

www.thesgc.org

Dr. Harding is the recipient of the Huntington's Disease Society of America Berman Topper Career Development Fellowship which funds and supports this research, in addition to generous funding from the CHDI Foundation and the Huntington Society of Canada. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome.

SGC

LAB SCRIBBLES
DR. RACHEL HARDING

labscribbles.com

zenodo

zenodo.org/communities/labscribbles

@labscribbles

www.thesgc.org

Dr. Harding is the recipient of the Huntington's Disease Society of America Berman Topper Career Development Fellowship which funds and supports this research, in addition to generous funding from the CHDI Foundation and the Huntington Society of Canada. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome.