Journal of Visualized Experiments

X-Ray Crystallography to Study the Oligomeric State Transition of the Thermotoga Maritima M42 Aminopeptidase TmPep1050

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE61288R1
Full Title:	X-Ray Crystallography to Study the Oligomeric State Transition of the Thermotoga Maritima M42 Aminopeptidase TmPep1050
Section/Category:	JoVE Biochemistry
Keywords:	Protein oligomerization; M42 aminopeptidase; oligomeric state transition; metalloenzyme; protein purification; protein crystallization; x-ray crystallography; data processing; molecular replacement
Corresponding Author:	Raphael Dutoit BELGIUM
Corresponding Author's Institution:	
Corresponding Author E-Mail:	Raphael.Dutoit@ulb.ac.be
Order of Authors:	Raphaël Dutoit
	Nathalie Brandt
	Dany Van Elder
	Louis Droogmans
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Brussels, Belgium

TITLE:

X-Ray Crystallography to Study the Oligomeric State Transition of the *Thermotoga Maritima* M42 Aminopeptidase TmPep1050

AUTHORS AND AFFILIATIONS:

Raphaël Dutoit[1,2], Nathalie Brandt[2], Dany Van Elder[1], Louis Droogmans[1]

[1]Laboratory of Microbiology, Department of Molecular Biology, Université Libre de Bruxelles, Brussels, Belgium

[2]Labiris Institut de Recherche, Brussels, Belgium

Corresponding Author:

Raphaël Dutoit (rdutoit@ulb.ac.be)

Email Addresses of Co-Authors:

Nathalie Brandt (nbrandt@spfb.brussels)

Dany Van Elder (dvelder@ulb.ac.be)

Louis Droogmans (ldroogma@ulb.ac.be)

KEYWORDS:

protein oligomerization, M42 aminopeptidase, oligomeric state transition, metalloenzyme, protein purification, protein crystallization, X-ray crystallography, data processing, molecular replacement

SUMMARY:

This protocol has been developed to study the dimer-dodecamer transition of TmPep1050, an M42 aminopeptidase, at the structural level. It is a straightforward pipeline starting from protein purification to X-ray data processing. Crystallogenesis, data set indexation, and molecular replacement have been emphasized through a case of study, $TmPep1050_{H60A\ H307A}$ variant.

ABSTRACT:

The M42 aminopeptidases form functionally active complexes made of 12 subunits. Their assembly process appears to be regulated by their metal ion cofactors triggering a dimer-dodecamer transition. Upon metal ion binding, several structural modifications occur in the active site and at the interaction interface, shaping dimers to promote the self-assembly. To observe such modifications, stable oligomers must be isolated prior to structural study. Reported here is a method that allows the purification of stable dodecamers and dimers of TmPep1050, an M42 aminopeptidase of *T. maritima*, and their structure determination by X-ray crystallography. Dimers were prepared from dodecamers by removing metal ions with a chelating agent. Without its cofactor, dodecamers became less stable and were fully dissociated upon heating. The oligomeric structures were solved by the straightforward molecular replacement approach. To illustrate the methodology, the structure of a TmPep1050 variant, totally impaired in metal ion binding, is presented showing no further breakdown of dimers to monomers.

INTRODUCTION:

Oligomerization is a predominant process that dictates the biological functions of many proteins. In *Escherichia coli*, it is estimated that only 35% of proteins are monomeric[1]. Some proteins, called morpheeins, can even adopt several oligomeric states with subunits having distinct structure in each oligomeric state[2]. The transition between their oligomeric states is often a mean to regulate protein activity as each oligomeric state may have a different specific activity or function. Several examples of morpheeins have been well-documented in literature, notably the porphobilinogen synthase[3], HPr kinase/phosphatase[4], Lon protease[5], lactate dehydrogenase[6], glyceraldehyde-3-phosphate dehydrogenase[7], pyruvate kinase[8], citrate synthase[9], and ribonuclease A[10]. Recently, we described the M42 aminopeptidase TmPep1050, another example of enzyme with morpheein-like behavior, whose activity depends on its oligomeric states[11]. The transition between its oligomeric states is mediated by its metallic cofactors which induce several structural modifications of the subunits.

The M42 aminopeptidase family belongs to the MH clan[12,13], and is widely distributed among Bacteria and Archaea[14]. The M42 aminopeptidases are genuine dinuclear enzymes degrading peptides up to 35 amino acid residues in length[15]. They adopt a peculiar tetrahedron-shaped structure made of 12 subunits with their active sites are oriented towards an inner cavity. Such an arrangement is often described as a nano-compartmentalization of the activity to avoid uncontrolled proteolysis. The physiological function of the M42 aminopeptidases may be associated with the proteasome, hydrolyzing peptides generated from protein degradation[16,17]. Indeed, *Pyrococcus horikoshii* possesses four M42 aminopeptidases, each presenting distinct but complementary specificities[18–21]. Singularly, heterocomplexes made of two different types of subunits have been described in *P. horikoshii*, suggesting the existence of peptidasome complexes[22,23].

Several structures of M42 aminopeptidases have been described in the literature[11,16,18–20,24–26]. The subunit is composed of two distinct domains, a catalytic domain and a dimerization domain. The catalytic domain adopts a common α/β domain conserved in the whole MH clan, the archetypal catalytic domain being the aminopeptidase Ap1 of *Vibrio proteolyticus*[27]. The dimerization domain adopts a PDZ-like fold[16] and may have, in addition to its role in the oligomerization, a role in controlling substrate access and binding in the inner cavity[11]. As the basic building block is a dimer, the dodecamer is often described as the association of six dimers, each dimer being positioned at each edge of the tetrahedron[16]. The oligomerization of the M42 aminopeptidases relies on the availability of its metal cofactors. Divalent metal ions, often Zn^{2+} and Co^{2+}, are catalytically involved in the peptide binding and hydrolysis. They are found in two distinct binding sites, namely M1 and M2 sites. The two metal ions also drive and finely tune the oligomerization as demonstrated for PhTET2, PhTET3, PfTET3, and TmPep1050[11,24,28,29]. When the metal cofactors are depleted, the dodecamer disassembles into dimers, like in PhTET2, PhTET3, and TmPep1050[11,16,28], or even monomers, like in PhTET2 and PfTET3[24,29].

Presented here is a protocol used for studying the structures of TmPep1050 oligomers. This protocol is a set of common methods including protein purification, proteolytic activity screening, crystallization, X-ray diffraction, and molecular replacement. Subtleties inherent to dealing with

metalloenzymes, protein oligomerization, protein crystallization and molecular replacement are emphasized. A case of study is also presented to show whether TmPep1050 dodecamers may further dissociate into monomers or not. To address this question, a TmPep1050 variant, $TmPep1050_{H60A\ H307A}$, has been studied whose metal binding sites are impaired by mutating His-60 (M2 site) and His-307 (M1 site) to Ala residues. This protocol may be accommodated to study other M42 aminopeptidases or any metalloenzymes with morpheein-like behavior.

PROTOCOL:

1. Production and purification of recombinant TmPep1050

NOTE: Hereafter are described the cloning procedure and purification of wild-type TmPep1050 adapted from a previous study[11]. Alternatively, the cloning can be done using a synthetic gene. To generate TmPep1050 variants, site directed mutagenesis can be performed following, for instance, the single-primer reactions in parallel protocol (SPRINP) method[30]. The purification protocol can be used for TmPep1050 variants. The use of His-tag should be avoided as it interferes with metal ion binding.

1.1. Expression vector design

1.1.1. Acquire genomic DNA of *Thermotoga maritima* MSB8 (ATCC 43589) or TmCD00089984 (Joint Center for Structural Genomics).

1.1.2. Amplify TM_1050 open reading frame (ORF) using either genomic DNA or template plasmid, a high-fidelity DNA polymerase, and the following primers: ocej419 (5'-TTTAACTTTAAGAAGGAGATATACATACCCATGAAGGAACTGATCAGAAAGCTG) and ocej420 (5'-ATCCGCCAAAACAGCCAAGCTGGAGACCGTTTACGCCCCCAGATACCTGATGAG). Run the polymerase chain reaction (PCR) screening according to the following scheme: 5 min at 95 °C, 30 cycles of 3 steps (30 s at 95 °C, 30 s at 55 °C, 90 s at 72 °C), and 10 min at 72 °C as the final step.

1.1.3. Clone the PCR fragment into a suitable expression vector (**Table of Materials**) by homologous recombination (**Figure 1**) in *E. coli* according to the SLiCE protocol[31]. To 50 ng of linearized vector, add PCR fragment in a 10:1 molar ratio of fragment to vector, 1 µL of PPY strain extract, 50 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 10 mM ATP, and 10 mM dithiothreitol (DTT) for a reaction volume of 10 µL. Incubate for 1 h at 37 °C.

1.1.4. Transform chemically competent *E. coli* XL1 blue strain (or any *rec*A$^-$ suitable strain) with 1 µL of recombination reaction. Plate the cells on LB medium containing 100 µg/mL ampicillin. Incubate the plates overnight at 37 °C.

1.1.5. Pick colonies on fresh LB plates with 100 µg/mL ampicillin. Incubate the plates at 37 °C for a least 8 h.

1.1.6. Screen for positive candidates by colony PCR using suitable primer pair (5'-

ATGCCATAGCATTTTTATCC and 5'- ATTTAATCTGTATCAGGC if using the recommended vector listed in **Table of Materials**). With a microtip end, scratch a picked colony and transfer the cells to 20 µL of reaction mix containing 0.5 µM of each primer and 10 µL of a commercial Taq DNA polymerase mix.

1.1.7. Run the PCR screening according to the following scheme: 5 min at 95 °C as the denaturation step, 30 cycles of 3 steps (30 s at 95 °C, 30 s at 55 °C, 90 s at 72 °C), and 10 min at 72 °C as the final step.

NOTE: The PCR reactions can be stored overnight in the PCR machine at 12 °C.

1.1.8. Load 10 µL of each PCR reaction on a 0.8% agarose gel prepared in Tris-acetate-EDTA (TAE) buffer. Run the electrophoresis for 25 min at 100 V.

NOTE: A 1.1 kbp amplicon is expected.

1.1.9. Extract plasmids from candidates using a commercial kit (**Table of Materials**) and sequence them using the same primer pair used in step 1.1.6.

1.2. Cell culture

NOTE: When a suitable candidate has been identified by sequencing, the clone can be directly used as the expression if using the recommended vector (**Table of Materials**). In that case, the expression is controlled by the arabinose-inducible P_{BAD} promoter[32].

1.2.1. Inoculate 10 mL of LB medium containing 100 µg/mL ampicillin with the candidate and incubate the preculture overnight at 37 °C under orbital shaking. Add 5 mL of the preculture to 1 L of LB medium with 100 µg/mL ampicillin. Mind to respect an air to liquid ratio of 3.

1.2.2. Let cells grow at 37 °C under orbital shaking. Monitor the optical density at 660 nm (OD_{660}).

1.2.3. When OD_{660} has reached 0.5−0.6, rapidly cool the culture for 5 min on ice and transfer it to an incubator set to 18 °C.

1.2.4. Add 0.2 g/L arabinose to induce gene expression and incubate for 12−18 h at 18 °C.

1.2.5. Harvest cells by centrifuging the culture at 6,000 x *g* for 30 min at 4 °C. Discard the supernatant and wash cells with 100 mL of 0.9% (w/v) NaCl.

1.2.6. Centrifuge again at 6,000 x *g* for 15 min at 4 °C and discard the supernatant.

NOTE: Cell pellets can be used directly for protein extraction or stored at -80 °C.

1.3. Protein purification

1.3.1. Resuspend the cell pellets in 40 mL of 50 mM MOPS, 1 mM $CoCl_2$, pH 7.2. Add 1 µL of 25 U/µL DNA/RNA endonuclease and one tablet of protease inhibitor cocktail that does not contain EDTA. Sonicate the suspension in pulse mode under cooling for 30 min.

1.3.2. Centrifuge the crude extract at 20,000 x *g* for 30 min at 4 °C. Collect the supernatant and heat it in a water bath at 70 °C for 10 min.

1.3.3. Centrifuge the denatured cell extract at 20,000 x *g* for 30 min at 4 °C and collect the supernatant for purification.

1.3.4. Use suitable anion exchange resin (**Table of Materials**) packed in a column of ~15 mL of volume. Refer to manufacturer's recommendations for the working flow rate and column pressure limit. Equilibrate the resin with 50 mM MOPS, 1 mM $CoCl_2$, pH 7.2.

1.3.5. Load the supernatant collected from step 1.3.3 into the column. Monitor the absorbance of the eluate at 280 nm. When it has reached the baseline, proceed to the elution.

1.3.6. Apply a gradient from 0 to 0.5 M NaCl in 50 mM MOPS, 1 mM $CoCl_2$, pH 7.2 for 5 column volumes (CV). Wait till the conductivity is stabilized and the absorbance has reached the baseline.

1.3.7. Apply a final gradient from 0.5 to 1 M NaCl in 50 mM MOPS, 1 mM $CoCl_2$, pH 7.2 for 1 CV.

1.3.8. Analyze some fractions (see **Figure 2A** for guidance) by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

NOTE: TmPep1050 appears as a 36 kDa band after Coomassie staining. Alternatively, the presence of TmPep1050 can be confirmed by activity assay (see section 2.1). At this step, fractions can be stored at 4 °C overnight.

1.3.9. Pool the fractions containing TmPep1050 and add finely ground powder of $(NH_4)_2SO_4$ to obtain a concentration of 1.5 M $(NH_4)_2SO_4$. Mix gently by inverting the tube upside down till complete dissolution.

1.3.10. Use hydrophobic interaction resin (**Table of Materials**) packed in a column of ~30 mL of volume. Refer to manufacturer's recommendations for the working flow rate and column pressure limit. Equilibrate the resin with 50 mM MOPS, 1.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$, pH 7.2.

1.3.11. Load the sample into the column and monitor the absorbance of the eluate at 280 nm. When the absorbance has reached the baseline, elute bound proteins by applying a gradient from 1.5 M to 0 M $(NH_4)_2SO_4$ in 50 mM MOPS, 1 mM $CoCl_2$, pH 7.2 for 5 CV.

1.3.12. Analyze some fractions (see **Figure 2B** for guidance) by SDS-PAGE.

NOTE: TmPep1050 appears as a 36 kDa band after Coomassie staining. Alternatively, the presence of TmPep1050 can be confirmed by activity assay (see section 2.1). At this step, fractions can be stored at 4 °C overnight.

1.3.13. Pool the fractions containing TmPep1050 and concentrate to 2 mL using ultrafiltration units with 30 kDa cutoff (**Table of Materials**). Proceed to section 1.4 to determine the molecular weight.

1.4. Size exclusion chromatography

1.4.1. Use size exclusion resin (**Table of Materials**) packed in a column of ~120 mL of volume. Refer to manufacturer's recommendations for the working flow rate and column pressure limit. Equilibrate the resin with 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$, pH 7.2.

1.4.2. Load the sample into the column and monitor the absorbance of the eluate at 280 nm. Fractionate from the column dead volume (~0.33 CV) till the end of the elution (1 CV).

1.4.3. Measure the elution volume for each observed peak.

NOTE: For guidance, dodecameric TmPep1050 elutes at ~82 mL (**Figure 3A**) under current experimental conditions, while dimeric TmPep1050, such as the $TmPep1050_{H60A\ H307A}$ variant, elutes at ~95 mL (**Figure 3B**). Some TmPep1050 may adopt both oligomeric forms, such as $TmPep1050_{H60A}$ (**Figure 3C**).

1.4.4. Analyze fractions corresponding to the maxima and tails of observed peaks using SDS-PAGE.

NOTE: TmPep1050 appears as a 36 kDa band after Coomassie staining.

1.4.5. Pool the fractions of each peak and concentrate using ultrafiltration units with 30 kDa cutoff (**Table of Materials**) to obtain a concentration of ~300 µM.

1.4.6. Measure the absorbance at 280 nm on a nano-volume spectrophotometer and calculate the concentration using the molecular extinction coefficient of 18,910 M^{-1} cm^{-1}.

1.4.7. Store the purified protein at -18 °C.

1.4.8. To determine the molecular weight, calibrate the size exclusion chromatography (SEC) column using molecular weight standards (**Table of Materials**). Analyze the standards using 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, 1 mM $CoCl_2$ pH 7.2 as the running buffer.

2. Activity assay and apo-enzyme preparation

NOTE: Originally, the apo-enzyme was prepared by diluting 1 volume of TmPep1050 in 10

volumes of 2.1 M malic acid pH 7.0 and concentrated back to 1 volume prior to dialysis[11]. Below is presented an alternative procedure using 1,10-phenanthroline, a metal ion chelator. This procedure reduces protein loss and gives the same results than the previously published method.

2.1. Activity assay

2.1.1. Prepare a stock solution of 100 mM L-Leucine-*p*-nitroanilide (**Table of Materials**) in methanol.

2.1.2. Add 25 µL of 100 mM L-Leucine-*p*-nitroanilide in 965 µL of 50 mM MOPS, 250 µM $CoCl_2$, pH 7.2, 10% methanol. Preincubate the reaction mix at 75 °C in a dry bath.

2.1.3. Dilute the enzyme in 50 mM MOPS pH 7.2 to a concentration of 1 µM. Add 10 µL to the reaction mix, vortex, and incubate at 75 °C either until it has turned yellowish or for 1 h.

2.1.4. Stop the reaction by adding 1 mL of 20% acetic acid. Vortex well and let it cool down to room temperature.

2.1.5. Transfer the reaction mix in a spectrophotometer cell. Read the absorbance at 410 nm against a negative control (incubated reaction mix without enzyme).

2.2. Apo-enzyme preparation

2.2.1. Prepare a stock solution of 1 M 1,10-phenanthroline in ethanol. Add 10 µL of 1,10-phenanthroline stock solution to 890 µL of 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, pH 7.2. Add 100 µL of purified TmPep1050 (300 µM−1 mM concentration).

2.2.2. Check the activity loss using the activity assay described in section 2.1 without adding $CoCl_2$ in the reaction mix.

2.2.3. Transfer the sample in a dialysis tube. Dialyze against 200 mL of 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, pH 7.2 at 4 °C. Exchange thrice the dialysate with fresh buffer during the 48 h dialysis.

2.2.4. Collect the sample from the dialysis tube and concentrate back to 100 µL using ultrafiltration units with 30 kDa cutoff (**Table of Materials**). Check the concentration by reading the absorbance at 280 nm using a nano-volume spectrophotometer.

2.3. Dimer preparation

2.3.1. Dilute the apo-enzyme to a concentration of 1 µM in 50 mM MOPS, 0.5 M $(NH_4)_2SO_4$, pH 7.2. Incubate for 2 h at 75 °C in a dry bath, then let the sample cool down to room temperature.

2.3.2. Concentrate the sample to an enzyme concentration of at least 50 µM. Check the molecular weight by SEC (see section 1.4). The elution peak must shift from ~82 mL to ~95 mL

(under current experimental conditions).

3. TmPep1050 crystallization

NOTE: Protein crystallization remains an empirical science as it is a multifactorial phenomenon[33]. While some parameters can be identified and controlled (such as temperature, pH, precipitation agent concentration), others may influence elusively the crystallization (such as protein and chemical purity, proteolysis, sample history). Nowadays protein crystallization is tackled in a rational and systematic manner thanks to a bunch of commercial crystallization screening conditions and automation. The optimization of a crystallization condition, however, relies mostly on a trial-and-error approach. Hereafter are described a blueprint for crystallizing proteins and several tips for optimizing the crystallization conditions.

3.1. Crystallization screening

NOTE: Using commercial crystallization kits, crystals of dodecameric TmPep1050 have been obtained in 2.2 M DL-malic acid pH 7.0, 0.1 M Bis-Tris propane pH 7.0 and 0.18 M tri-ammonium citrate, 20% polyethylene glycol (PEG) 3350. Crystals of dimeric TmPep1050 have been obtained in 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate, 30% PEG4000. Crystals of dodecamers appear within a week and reach their full size in a month. Crystals of dimers usually appear within 24 h and grow to full size in a week.

3.1.1. Acquire several commercial crystallization kits (see **Table of Materials** for examples).

3.1.2. Set up crystallization plates (**Table of Materials**) for the hanging drop method. Fill the wells with 500 µL of each solution of a crystallization screening kit.

3.1.3. For each well, set up a crystallization support. On the support, deposit a 1 µL drop of purified protein (usually ~10 mg/mL).

3.1.4. Immediately pipet 1 µL of crystallization solution from the well. Add it carefully to the protein drop and mix gently by pipetting upside down thrice. The drop must remain semispherical without any bubbles.

3.1.5. Screw the support on top of the corresponding well. Repeat the operation for the whole kit.

3.1.6. After setting up the plates, observe each drop with a binocular. Refer to the crystallization kit user guide for interpretation (clear drop, phase separation, precipitate, needles, etc.).

3.1.7. Incubate the plates at 20 °C. Check the plates once per day during the first week and once per week afterwards.

3.1.8. Score each well using the score sheet and the user guide provided with the crystallization

kits.

3.2. Crystallization optimization

NOTE: The initial crystallization conditions of dodecameric TmPep1050 have been optimized to 2.1 M DL-malic acid pH 6.75 and 0.18 M tri-ammonium citrate pH 7.5, 40% (w/v) PEG3350 while the crystallization condition of dimeric TmPep1050 has been shifted to 0.1 M sodium citrate pH 6.0, 10% (w/v) PEG3350. One cycle of seeding has been necessary to improve crystallinity. Hereafter is described how the crystallization of $TmPep1050_{H60A\ H307A}$ variant has been optimized.

3.2.1. Prepare stock solutions of 0.5 M sodium citrate buffer at different pH (4.5, 5.2, and 6.0), and 50% (w/v) PEG3350 solution.

3.2.2. Set up a crystallization plate as a matrix of pH vs. precipitation agent (see **Figure 4A**).

3.2.3. Incubate the plate at 20 °C. Observe each well with a binocular once per day during a week.

3.2.4. Score each well according to crystal size and shape. Select a condition giving crystals of at least 50 µm. Proceed to microseeding.

3.3. Microseeding

NOTE: Microseeding is a powerful method to improve the shape, size and crystallinity of protein crystals[34]. A faster seeding approach is streak seeding using a cat whisker. See **Figure 4** as an example how crystallization optimization and microseeding have improved the crystal shape and size for $TmPep1050_{H60A\ H307A}$.

3.3.1. Prepare the seeds using the selected well in step 3.2.4.

3.3.2. Increase the volume of a drop containing crystals to 10 µL by adding crystallization solution from the well. Pipet the drop and add 90 µL of crystallization solution from the well. Vortex thoroughly and keep the seeds on ice.

3.3.3. Prepare several dilutions of the seeds: 1x, 10x, 25x, and 100x. Vortex well the seeds before pipetting. Keep the dilutions on ice.

3.3.4. For each seed dilution, set up a crystallization plate as a matrix of pH vs. precipitation agent (see **Figure 4B**). Use the stock solutions prepared in step 3.2.1.

3.3.5. When making the drop, add 0.2 µL of seeds for a 2 µL drop. Incubate the plate at 20 °C. Observe each well with a binocular once per day during a week.

NOTE: Crystal size distribution and shape must be improved, see **Figure 4C** as an example for $TmPep1050_{H60A\ H307A}$. For further uses, seeds can be stored at -20 °C.

4. X-ray diffraction

4.1. Crystal picking

NOTE: Sample preparation depends on the X-ray source facility (home facility vs. synchrotron). Use storage devices (vials and vial-holder basket) accordingly. The addition of cryoprotectant (such as glycerol) may be required depending on salt/precipitation agent concentration. For TmPep1050 crystals, a cryoprotectant is not necessary as PEG or buffer concentration is high enough to avoid water crystals.

4.1.1. Prepare a bath filled with liquid nitrogen, plunge any vials or basket used for sample handling.

4.1.2. Set up sample picking loops of different sizes: 100, 150, and 200 µm (**Table of Materials**). Choose the sizes according to the crystal size.

4.1.3. Using a binocular, check the drop containing crystals and spot isolated crystals (the easiest to pick). With a loop, gently pick a crystal from the bottom. Immediately plunge the loop in liquid nitrogen and place the loop in a suitable vial.

4.2. Data collection

NOTE: Data collection may vary greatly depending on the X-ray source (home facility vs. synchrotron) and detector sensitivity. The collection strategy may also differ greatly from a sample to another, depending on the resolution, spot intensity, space group, etc. The topic has been extensively reviewed by Dauter[35].

4.2.1. Mount the loop carrying the crystal on the goniometer head of the diffractometer.

4.2.2. Tune the goniometer head along XYZ axes to align the crystal with the X-ray beam path.

4.2.3. Set the wavelength to 0.98 Å and move the detector to get 2 Å of resolution.

4.2.4. Start a short data collection by acquiring images in at least two different crystal orientations. Take 10 images (1 image per 0.1°) at 0° and 90°.

4.2.5. Check the collected images with suitable software (e.g., ADXV, XDS-Viewer or Albula Viewer). Determine the spot intensity and the highest resolution where spots are seen. Check also the monocrystallinity and spot separation.

4.2.6. Eventually, repeat steps 4.2.3–4.2.5 by changing the detector position for higher or lower resolution and the exposure time in accordance to the observation.

4.2.7. Start data collection around 360° with 1 image taken per 0.1°. Remember to set the detector position and exposure time optimally.

5. Indexation, molecular replacement and model building

5.1. Indexation

NOTE: Indexation is a method for measuring diffraction spots intensity, giving the amplitudes of structure factors[36]. Four software packages are commonly used for processing collected images: Mosflm[37], HKL2000[38], DIALS[39], and XDS[40]. The latter has been used for indexing the data sets obtained from TmPep1050 crystal diffraction.

5.1.1. Install XDS package and XDSME[41]. If processing HDF5 files, install XDS Neggia plugin (available on Dectris website). For more information, visit XDS wiki web page https://strucbio.biologie.uni-konstanz.de/xdswiki/index.php/Main_Page and XDSME web page https://github.com/legrandp/xdsme.

5.1.2. Before processing data, create a folder from where XDS will be run. Locate the path to the images.

5.1.3. To run XDSME, type **xdsme /path_to_images/image.extension** in a terminal window.

5.1.4. After XDS has ended the job, check the CORRECT.LP file. Note the probability of the space group determination, data completeness, the highest resolution, crystal mosaicity, and data quality. Check also XDS_pointless.log to obtain the likelihood of space groups.

NOTE: See **Figure 5** as an example of output.

5.1.5. Rerun XDSME with different space group solutions proposed by XDS in separate folder to avoid overwriting the previous process. Type **xdsme -s space_group_name -c "unit_cell_parameters" /path_to_images/image.extension** (e.g., **xdsme -s P21 -c "43.295 137.812 61.118 90.000 110.716 90.000**).

5.1.6. Check the CORRECT.LP files and choose the best solution based on data statistics.

5.1.7. Run XSCALE by typing **xscale.py XDS_ASCII.HKL**. Run XDSCONV by typing **xdsconv.py XSCALE.HKL ccp4**.

NOTE: In some cases, XDSME fails to identify the space group or fails to cut the resolution range properly or generates weird data statistics. If such a problem is encountered, it is worth to run XDS natively. Several parameters must be introduced in the XDS.INP initiation file (see XDS wiki page). When using XDS, the likelihood of possible space groups can be checked by using Pointless, part of CCP4 package[42]. To cut the data set resolution, R_{meas} < 60% and I/σ ~2 are commonly accepted to determine the highest resolution[43]. The molecular replacement and model

refinement, however, can be improved by extending the resolution to I/σ ~0.5−1.5 and $CC_{1/2}$ down to 0.2−0.4[44].

5.2. Molecular replacement

NOTE: Experimental data give access to the amplitude of structural factors but, without knowing the phase, they are useless. The phase can be determined experimentally by different methods relying on an anomalous signal (from a heavy atom, for instance)[45]. Molecular replacement is another method for determining the phase without an anomalous scattering atom[46,47]. This method uses the coordinates of a related molecule to find and improve the phase iteratively. We use Phaser[48] in Phenix GUI[49] for molecular replacement.

5.2.1. Prepare the starting model for molecular replacement using 4P6Y coordinates. From the pdb file, extract the monomer A and truncate its aminoacids in alanine using the PDB file editor in Phenix (under the **Model tools** tab).

5.2.2. Run Xtriage in Phenix (under the **Data analysis** tab) with the reflection file generated by XDSCONV (5.1.9) and the sequence as inputs.

5.2.3. Check the log file from Xtriage. Note the completeness, the number of subunits in the asymmetric unit, the anisotropy, the presence of ice rings, and twinning occurrence.

5.2.4. Run Phaser-MR in Phenix (under the **Molecular replacement** tab) for molecular replacement using the reflection file, the sequence and the starting 4P6Y model truncated in poly alanine (step 5.2.1).

5.2.5. Upon completion, check if a model has been found and the score of the molecular replacement. A translation factor Z-score (TFZ) of at least 8 indicates that the solution is definitively correct.

5.3. Model building

NOTE: After determining the phase by molecular replacement, the model must be built and refined. This protocol uses Phenix GUI[49] for automatic building and iterative refinements, and Coot[50] for manual structure building and refinement.

5.3.1. After molecular replacement using Phase-MR in Phenix, select **Run Autobuild**. All the required files will be automatically added. Simply press **Run** to start autobuild.

5.3.2. Upon completion, check the model in Coot. Build and refine the model manually according the electron density map in Coot.

5.3.3. Refine the manually curated model in Phenix (in the **Refinement** tab) using the model, the sequence, and the diffraction data as inputs. Refer to Phenix help to choose the right strategy.

5.3.4. After refinement, check the results: R_{free} and R_{work} must decrease, Molprobity[51] indicators must be respected, and outliers with low real-space correlation must be limited.

5.3.5. Repeat steps 5.3.2−5.3.4 until the best refined model is generated.

5.3.6. Run Molprobity on the server: http://molprobity.biochem.duke.edu/. Check any outliers identified by Molprobity.

5.3.7. Eventually repeat steps 5.3.2−5.3.6 until the best refined model is obtained.

REPRESENTATIVE RESULTS:

To study a possible dodecamer dissociation into monomers in TmPep1050, the His-60 and His-307 codons were replaced by alanine codon using a synthetic gene. This gene was then cloned in pBAD vector for expression and purification of the corresponding TmPep1050 variant subsequently named TmPep1050$_{H60A\ H307A}$. Size exclusion chromatography (**Figure 3B**) showed that the purified protein had an apparent molecular weight of 56 kDa (molecular weight of the monomer being 36.0 kDa). A similar apparent molecular weight, 52 kDa, has been reported for TmPep1050 dimer[11]. Hence, the oligomeric state of TmPep1050$_{H60A\ H307A}$ could be inferred as dimeric. Regarding its specific activity, TmPep1050$_{H60A\ H307A}$ was completely inactive on L-Leu-*p*NA as substrate, even in the presence of cobalt ions. This result strongly suggests that the variants cannot bind any metal ions.

The crystallization condition of TmPep1050$_{H60A\ H307A}$ was optimized by varying pH vs. PEG concentration (**Figure 4**) around the condition of the dimer (i.e., 0.1 M sodium citrate pH 6.0 10% PEG3350). The best crystals of TmPep1050$_{H60A\ H307A}$ were obtained in 0.1 M sodium citrate pH 5.2 20% PEG3350 with one cycle of microseeding for improving monocrystallinity. A complete data set was collected at Proxima 2 beamline (SOLEIL synchrotron) at a resolution 2.36 Å (**Table 1**). Data indexation showed that the space group of the TmPep1050$_{H60A\ H307A}$ crystal is $C222_1$ but XDS proposed another solution, the *m*P space group (see **Figure 5**). According to Pointless, the likelihood of $C222_1$ and $P2_1$ space groups were 0.711 and 0.149, respectively. According to the data quality analysis, two monomers are found in the asymmetric unit. The analysis by Xtriage revealed that the data set is probably twinned but twinning in $C222_1$ space group is unlikely[52]. Twinning results from crystal growth anomaly where several definite domains have some of their lattice directions parallel to each other[53]. Twinning may also result from a higher crystal symmetry, indicating an erroneous data indexation. Hence, a pseudo-merohedral twin may exist so that a $P2_1$ crystal lattice looks like a $C222_1$. The data set was subsequently indexed in space group $P2_1$ and tested in molecular replacement. Xtriage analysis of the data set indexed in $P2_1$ revealed a pseudo-merohedral twin following a twin law *h, -k, -h-l*.

Using the coordinates of a monomer from dodecameric TmPep1050 (PDB code 4P6Y), a molecular replacement solution was found for the data set indexed in $P2_1$ only, with a TFZ score of 28.9. Therefore, the diffraction data were treated as a twinned data set for model building. To minimize the bias of molecular replacement, a first model was built by using

phenix.autobuild[54,55]. The structure of TmPep1050$_{H60A\ H307A}$ was completed after several cycles of automated and manual refinement in Phenix and Coot (**Table 1** and **Figure 6A**). The structure confirmed the oligomeric state with an interface surface of 1,710 Å^2 between both monomers and a Δ^iG of -16.2 kcal mol^{-1} as calculated by PDBe Pisa[56]. In comparison, the interface surface and Δ^iG of dimeric TmPep1050$_{2\text{-mer}}$ is 1,673 Å^2 and -16.7 kcal mol^{-1}, respectively.

The structure of TmPep1050$_{H60A\ H307A}$ is highly similar to the wild-type dimer structure with RMS of 0.774 Å upon alignment. Importantly, the same structural modifications are observed in both structure: high flexibility of the α8 and α10 helices, disordered active site Gln-196–Val-202, and the displacement of Lys-229–Ala-235 and Lys-247–Ser 254. These modifications were correlated previously with the hindrance of dodecamer formation in absence of its metal cofactor[11]. The two mutations of His-60 and His-307, however, had a slight effect on the side chains of Asp-168 and Asp-62. They appeared to be locked in a conformation different from the wild-type dimer (**Figure 6B**). The Asp-168 carboxylate is rotated by 40° due to the absence of His-60 and His-307. Hence both histidine residues are important for positioning the Asp-168 carboxylate correctly for bridging the two metal ions. The Asp-62 side chain is oriented towards Glu-18 carboxylate, outside the catalytic site. Asp-62 may have an important role in catalysis as it is assumed to modulate the pK_a of His-60 and, thus, influence metal ion binding in the M2 site. In addition, it could be implicated structurally in the stabilization of the catalytic site upon metal ion binding, favoring the formation of the dodecamer.

FIGURE AND TABLE LEGENDS:

Figure 1: Schematic representation of TM_1050 ORF cloning into pBAD vector by homologous recombination. The ORF is flanked by two 30 bp sequences homologous to the promoter BAD end and the sequence upstream *Pme*I restriction site.

Figure 2: Chromatograms of TmPep1050 purification. (**A**) Anion exchange chromatography. (**B**) Hydrophobic interaction chromatography. The absorbance (Abs), expressed in milliunits of absorbance (mUA), is shown in plain line. The conductivity, expressed in mS cm^{-1}, is shown in dashed line. The grey box indicates where TmPep1050 eluates on the chromatograms.

Figure 3: Size exclusion chromatography of (A) TmPep1050 dodecamer, (B) TmPep1050$_{H60A\ H307A}$, and (C) TmPep1050$_{H60A}$. Samples were analyzed using SEC resin packed in a 120 mL column. Absorbance (Abs) is expressed in milliunits of absorbance (mUA). (**D**) Calibration of the SEC column using thyroglobulin (T), ferritin (F), aldolase (Ald), conalbumin (C), and albumin (Alb) as standards. The correlation between the logarithm of the relative mass and the elution volume is linear, with a R^2 of 0.91. The 95% confidence intervals are represented as dots.

Figure 4: Optimization of TmPep1050$_{H60A\ H307A}$ crystallization. (**A**) The first optimization strategy consists of varying pH (between 4.5 and 6.0) vs. PEG3350 concentration (between 5% and 25%). The crystallization plate is schematized, and the wells are color coded: red for precipitate, yellow for polycrystals, and green for monocrystals. (**B**) The second optimization strategy includes the use of seeds diluted 25x with a narrower variation of pH vs. PEG3350. (**C**) Crystal shape and size before (upper image) and after (lower image) crystallization optimization and microseeding.

Figure 5: Excerpts from the log output CORRECT.LP of TmPep1050$_{H60A\ H307A}$ data indexation by XDS. Upper panel, the possible Bravais lattices, the most likely being *m*C, *m*P, and *o*C. Middle panel, overall statistics of data indexed in $C222_1$ space group. Lower panel, overall statistics of data indexed in $P2_1$ space group.

Figure 6: Structure of TmPep1050$_{H60A\ H307A}$. (**A**) Structural alignment of a TmPep1050$_{H60A\ H307A}$ subunit (red, PDB code 5NE9) vs. a dodecamer subunit (white, PDB code 6NW5) and a dimer subunit (blue, PDB code 5NE6). Arrows indicate the structural dissimilarities between dodecamers and dimers. (**B**) Close-up of the TmPep1050$_{H60A\ H307A}$ active site (red) compared to the active site of TmPep1050 dimer (blue) and dodecamer (white).

Table 1: Data collection and refinement statistics. Values in parentheses are for the highest-resolution shell.

DISCUSSION:

The protocol described herein allows understanding the dimer-dodecamer transition of TmPep1050 at the structural level. The methodology was experienced previously for determining the structure of both TmPep1050 oligomers[11]. The most challenging step was to find conditions promoting the dissociation of dodecamers into stable dimers. Such conditions had to be mild enough to permit the reassociation of dimers into dodecamers when the metal ion cofactor was added. The separation of oligomers was also a critical step as it conditions the structural studies and further biochemical characterization (e.g., studying the dodecamer reassociation in various doses of Co^{2+}). The molecular replacement, a proven method for phase determination, was used to solve the structures of TmPep1050 oligomers and its variants. The proposed protocol may be adapted to study other metallo-enzymes whose oligomerization states depend on the availability of their metal cofactors.

To illustrate the protocol, a case of study was presented, TmPep1050$_{H60A\ H307A}$ whose metal binding sites were impaired by mutating His-60 and His-307 to alanine. These residues bind Co^{2+} at the M2 and M1 sites, respectively. Interfering in metal binding could have perturbed the oligomerization state and led to a complete dissociation into monomers. Evidences of such a phenomenon have been reported for PhTET2 and PfTET3, two M42 aminopeptidases from *P. horikoshii* and *P. furiosus*, respectively[24,29]. TmPep1050$_{H60A\ H307A}$ did not behave as expected as this variant formed dimers only. Its structure showed the same modifications as the wild-type dimer but with two small exceptions. Indeed, the side chains of Asp-168 and Asp-62 appeared to be locked in an unconventional orientation preventing the stabilization of the active site. Their orientation seemed to be imposed by His-60 and His-307 as such modifications were not observed in the single point mutation variants.

ACKNOWLEDGMENTS:

We thank Martine Roovers for proofreading this paper and giving constructive comments. Access to Proxima 2 beamline (SOLEIL synchrotron) was within Block Allocation Groups 20151139.

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1. Levy, E. D., Teichmann, S. A. Structural, Evolutionary, and Assembly Principles of Protein Oligomerization. *Progress in Molecular Biology* and *Translational Science*. **117**, 25–51 (2013).
2. Selwood, T., Jaffe, E. K. Dynamic dissociating homo-oligomers and the control of protein function. *Archives of Biochemistry* and *Biophysics*. **519** (2), 131–143 (2012).
3. Jaffe, E. K. The Remarkable Character of Porphobilinogen Synthase. *Accounts of Chemical Research*. **49** (11), 2509–2517 (2016).
4. Ramström, H. et al. Properties and Regulation of the Bifunctional Enzyme HPr Kinase/Phosphatase in *Bacillus subtilis*. *Journal of Biological Chemistry*. **278** (2), 1174–1185 (2003).
5. Rudyak, S. G., Brenowitz, M., Shrader, T. E. Mg^{2+}-Linked Oligomerization Modulates the Catalytic Activiy of the Lon (La) Protease from *Mycobacterium smegmatis*. *Biochemistry*. **40** (31), 9317–9323 (2001).
6. Yamamoto, S., Storey, K. B. Dissociation-Association of lactate dehydrogenase Isozymes: Influences on the formation of tetramers vs. dimers of M4-LDH and H4-LDH. *International Journal of Biochemistry*. **20** (11), 1261–1265 (1988).
7. Sirover, M. A. Structural analysis of glyceraldehyde-3-phosphate dehydrogenase functional diversity. *The International Journal of Biochemistry & Cell Biology*. **57**, 20–26 (2014).
8. Gupta, V., Bamezai, R. N. K. Human pyruvate kinase M2: A multifunctional protein: Multifunctional Human PKM2. *Protein Science*. **19** (11), 2031–2044 (2010).
9. Wiegand, G., Remington, S. J. Citrate synthase: Structure, Control, and Mechanism. *Annual Review of Biophysics and Biophysical Chemistry*. **15**, 97–117 (1986).
10. Libonati, M., Gotte, G. Oligomerization of bovine ribonuclease A: structural and functional features of its multimers. *Biochemical Journal*. **380** (2), 311–327 (2004).
11. Dutoit, R. et al. How metal cofactors drive dimer–dodecamer transition of the M42 aminopeptidase TmPep1050 of *Thermotoga maritima*. *Journal of Biological Chemistry*. **294** (47), 17777–17789 (2019).
12. Rawlings, N. D. et al. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*. **46** (D1), D624–D632 (2018).
13. Neuwald, A. F., Liu, J. S., Lipman, D. J., Lawrence, C. E. Extracting protein alignment models from the sequence database. *Nucleic Acids Research*. **25** (9), 1665–1677 (1997).
14. Dutoit, R., Brandt, N., Legrain, C., Bauvois, C. Functional Characterization of Two M42 Aminopeptidases Erroneously Annotated as Cellulases. *PLoS ONE*. **7** (11), e50639 (2012).
15. Franzetti, B. et al. Tetrahedral aminopeptidase: a novel large protease complex from archaea. *The EMBO Journal*. **21** (9), 2132–2138 (2002).
16. Borissenko, L., Groll, M. Crystal Structure of TET Protease Reveals Complementary Protein Degradation Pathways in Prokaryotes. *Journal of Molecular Biology*. **346** (5), 1207–1219 (2005).
17. Appolaire, A. et al. TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. *Biochimie*. **122**, 188–196 (2016).

18. Russo, S., Baumann, U. Crystal Structure of a Dodecameric Tetrahedral-shaped Aminopeptidase. *Journal of Biological Chemistry*. **279** (49), 51275–51281 (2004).
19. Schoehn, G. et al. An Archaeal Peptidase Assembles into Two Different Quaternary Structures: *A tetrahedron* and *a giant octahedron*. *Journal of Biological Chemistry*. **281** (47), 36327–36337 (2006).
20. Durá, M. A. et al. The structural and biochemical characterizations of a novel TET peptidase complex from Pyrococcus horikoshii reveal an integrated peptide degradation system in hyperthermophilic Archaea: Characterization of P. horikoshii TET3 peptidase. *Molecular Microbiology*. **72** (1), 26–40 (2009).
21. Basbous, H., Appolaire, A., Girard, E., Franzetti, B. Characterization of a Glycyl-Specific TET Aminopeptidase Complex from *Pyrococcus horikoshii*. *Journal of Bacteriology*. **200** (17), e00059-18 (2018).
22. Appolaire, A. et al. Small-angle neutron scattering reveals the assembly mode and oligomeric architecture of TET, a large, dodecameric aminopeptidase. *Acta Crystallographica Section D Biological Crystallography*. **70** (11), 2983–2993 (2014).
23. Appolaire, A. et al. The TET2 and TET3 aminopeptidases from *P yrococcus horikoshii* form a hetero-subunit peptidasome with enhanced peptide destruction properties: TET aminopeptidase multi-subunit complex. *Molecular Microbiology*. **94** (4), 803–814 (2014).
24. Colombo, M., Girard, E., Franzetti, B. Tuned by metals: the TET peptidase activity is controlled by 3 metal binding sites. *Scientific Reports*. **6** (1), 20876 (2016).
25. Petrova, T. E. et al. Structure of the dodecamer of the aminopeptidase APDkam598 from the archaeon *Desulfurococcus kamchatkensis*. *Acta Crystallographica Section F Structural Biology Communications*. **71** (3), 277–285 (2015).
26. Kim, D. et al. Structural basis for the substrate specificity of PepA from Streptococcus pneumoniae, a dodecameric tetrahedral protease. *Biochemical* and *Biophysical Research Communications*. **391** (1), 431–436 (2010).
27. Chevrier, B. et al. Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure*. **2**, 283–291 (1994).
28. Rosenbaum, E., Ferruit, M., Durá, M. A., Franzetti, B. Studies on the parameters controlling the stability of the TET peptidase superstructure from Pyrococcus horikoshii revealed a crucial role of pH and catalytic metals in the oligomerization process. *Biochimica et Biophysica Acta (BBA) - Proteins* and *Proteomics*. **1814** (10), 1289–1294 (2011).
29. Macek, P. et al. Unraveling self-assembly pathways of the 468-kDa proteolytic machine TET2. *Science Advances*. **3** (4), e1601601 (2017).
30. Edelheit, O., Hanukoglu, A., Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology*. **9** (1), 61 (2009).
31. Zhang, Y., Werling, U., Edelmann, W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Research*. **40** (8), e55–e55 (2012).
32. Schleif, R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli* , and the light switch mechanism of AraC action. *FEMS Microbiology Reviews*. **34** (5), 779–796 (2010).
33. McPherson, A., Gavira, J. A. Introduction to protein crystallization. *Acta Crystallographica Section F Structural Biology Communications*. **70** (1), 2–20 (2014).
34. Bergfors, T. Seeds to crystals. *Journal of Structural Biology*. **142** (1), 66–76 (2003).

35. Dauter, Z. Collection of X-Ray Diffraction Data from Macromolecular Crystals. *Protein Crystallography*. **1607**, 165–184 (2017).
36. Powell, H. R. X-ray data processing. *Bioscience Reports*. **37** (5), BSR20170227 (2017).
37. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., Leslie, A. G. W. *iMOSFLM* : a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Crystallographica Section D Biological Crystallography*. **67** (4), 271–281 (2011).
38. Otwinowski, Z., Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*. **276**, 307–326 (1997).
39. Clabbers, M. T. B., Gruene, T., Parkhurst, J. M., Abrahams, J. P., Waterman, D. G. Electron diffraction data processing with *DIALS*. *Acta Crystallographica Section D Structural Biology*. **74** (6), 506–518 (2018).
40. Kabsch, W. XDS. *Acta Crystallographica Section D Biological Crystallography*. **66** (2), 125–132 (2010).
41. Legrand, P. *legrandp/xdsme: March 2019 version working with the latest XDS version (Jan 26, 2018)*. doi: 10.5281/ZENODO.837885. Zenodo. (2019).
42. Evans, P. Scaling and assessment of data quality. *Acta Crystallographica Section D Biological Crystallography*. **62** (1), 72–82 (2006).
43. Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures: Protein crystallography for non-crystallographers. *FEBS Journal*. **275** (1), 1–21 (2008).
44. Karplus, P. A., Diederichs, K. Assessing and maximizing data quality in macromolecular crystallography. *Current Opinion in Structural Biology*. **34**, 60–68 (2015).
45. Taylor, G. L. Introduction to phasing. *Acta Crystallographica Section D Biological Crystallography*. **66** (4), 325–338 (2010).
46. Rossmann, M. G., Blow, D. M. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica*. **15**, 24–31 (1962).
47. Rossmann, M. G. The molecular replacement method. *Acta Crystallographica Section A Foundations of Crystallography*. **46** (2), 73–82 (1990).
48. McCoy, A. J. et al. *Phaser* crystallographic software. *Journal of Applied Crystallography*. **40** (4), 658–674 (2007).
49. Adams, P. D. et al. *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*. **66** (2), 213–221 (2010).
50. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of *Coot*. *Acta Crystallographica Section D Biological Crystallography*. **66** (4), 486–501 (2010).
51. Chen, V. B. et al. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*. **66** (1), 12–21 (2010).
52. Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N., Adams, P. D. Surprises and pitfalls arising from (pseudo)symmetry. *Acta Crystallographica Section D Biological Crystallography*. **64** (1), 99–107 (2008).
53. Yeates, T. O. Detecting and overcoming crystal twinning. *Methods in Enzymology*. **276**, 344–358 (1997).

54. Terwilliger, T. C. Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Crystallographica Section D Biological Crystallography*. **60** (12), 2144–2149 (2004).
55. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard. *Acta Crystallographica Section D Biological Crystallography*. **64** (1), 61–69 (2008).
56. Krissinel, E., Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*. **372** (3), 774–797 (2007).

Figure 1

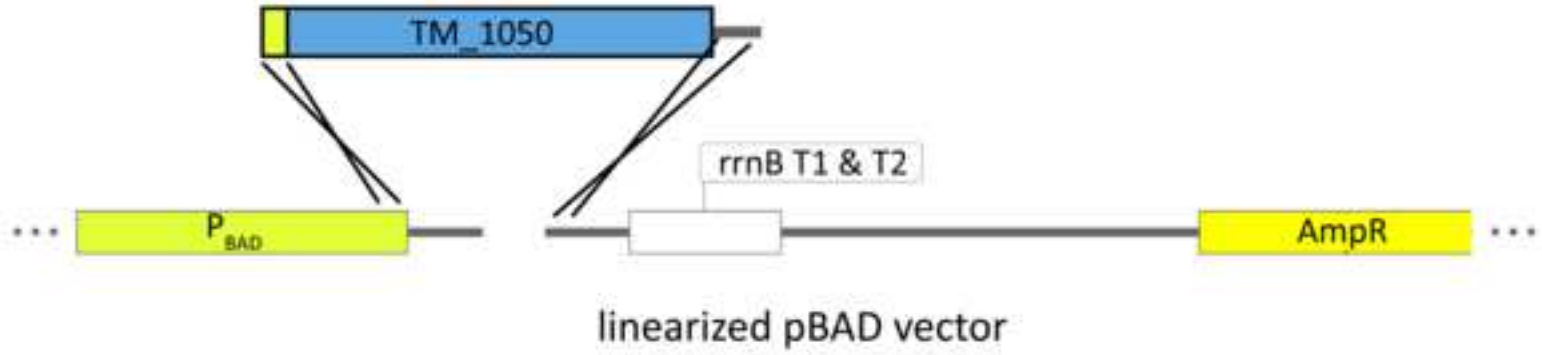

Figure 2

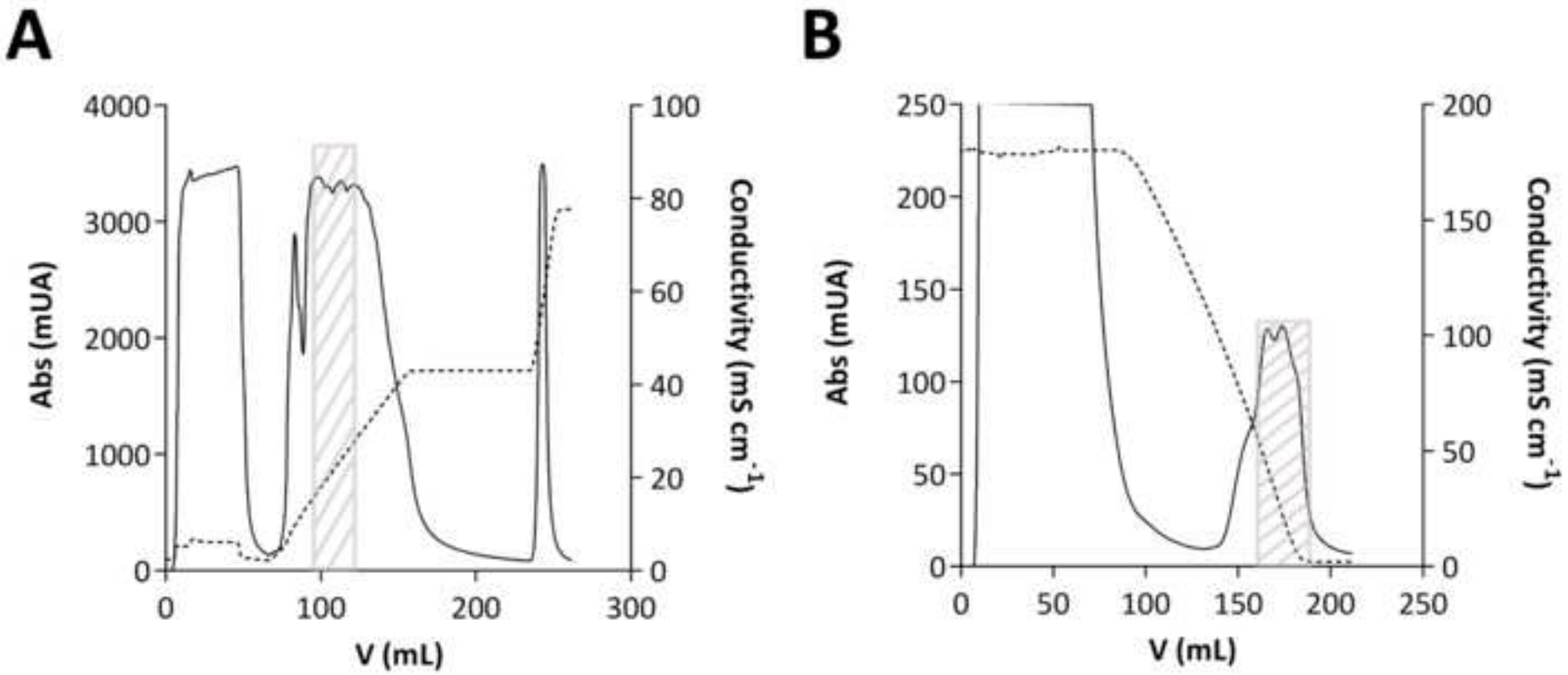

Figure 3

Click here to access/download;Figure;Figure3.psd

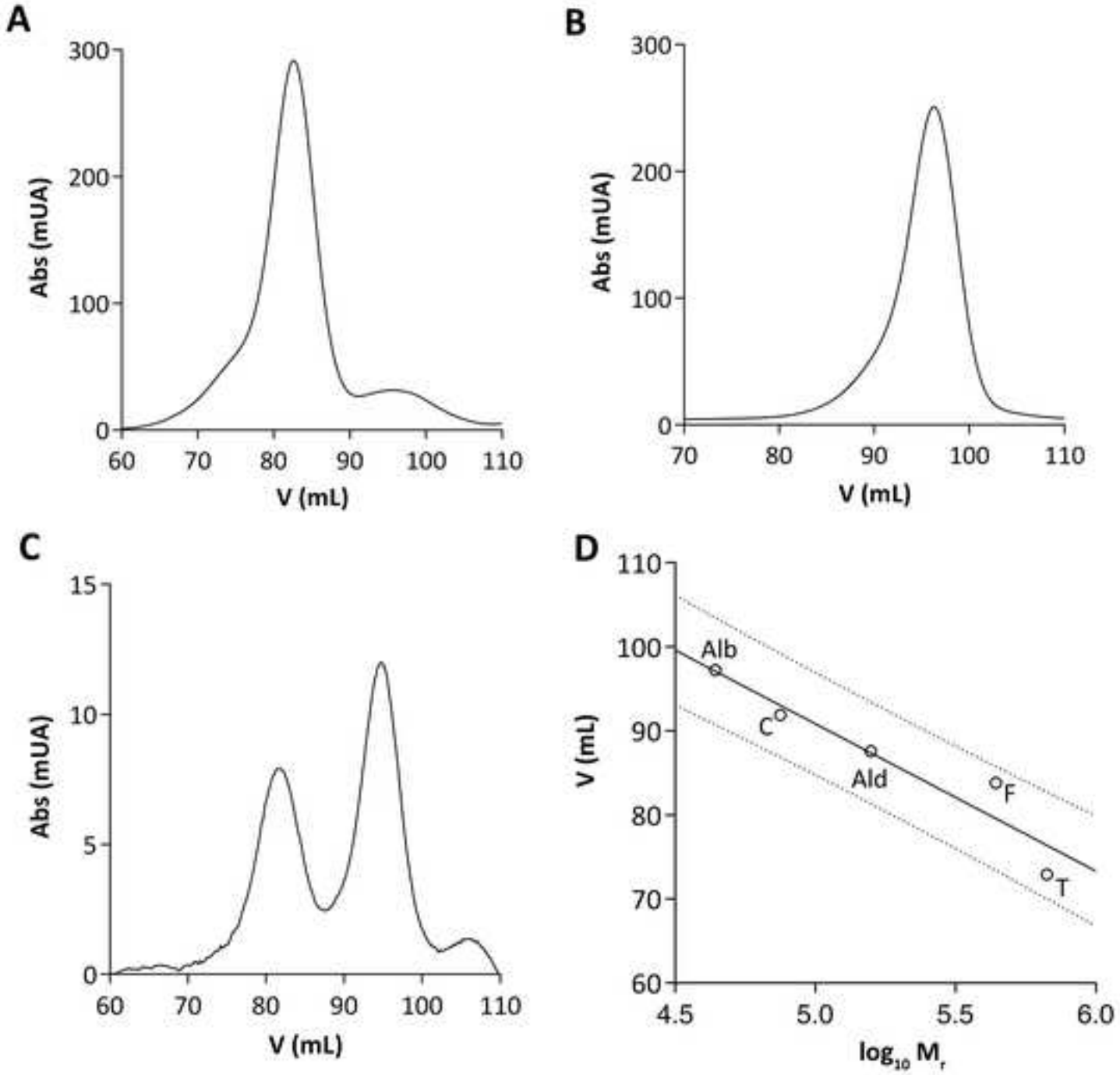

Figure 4

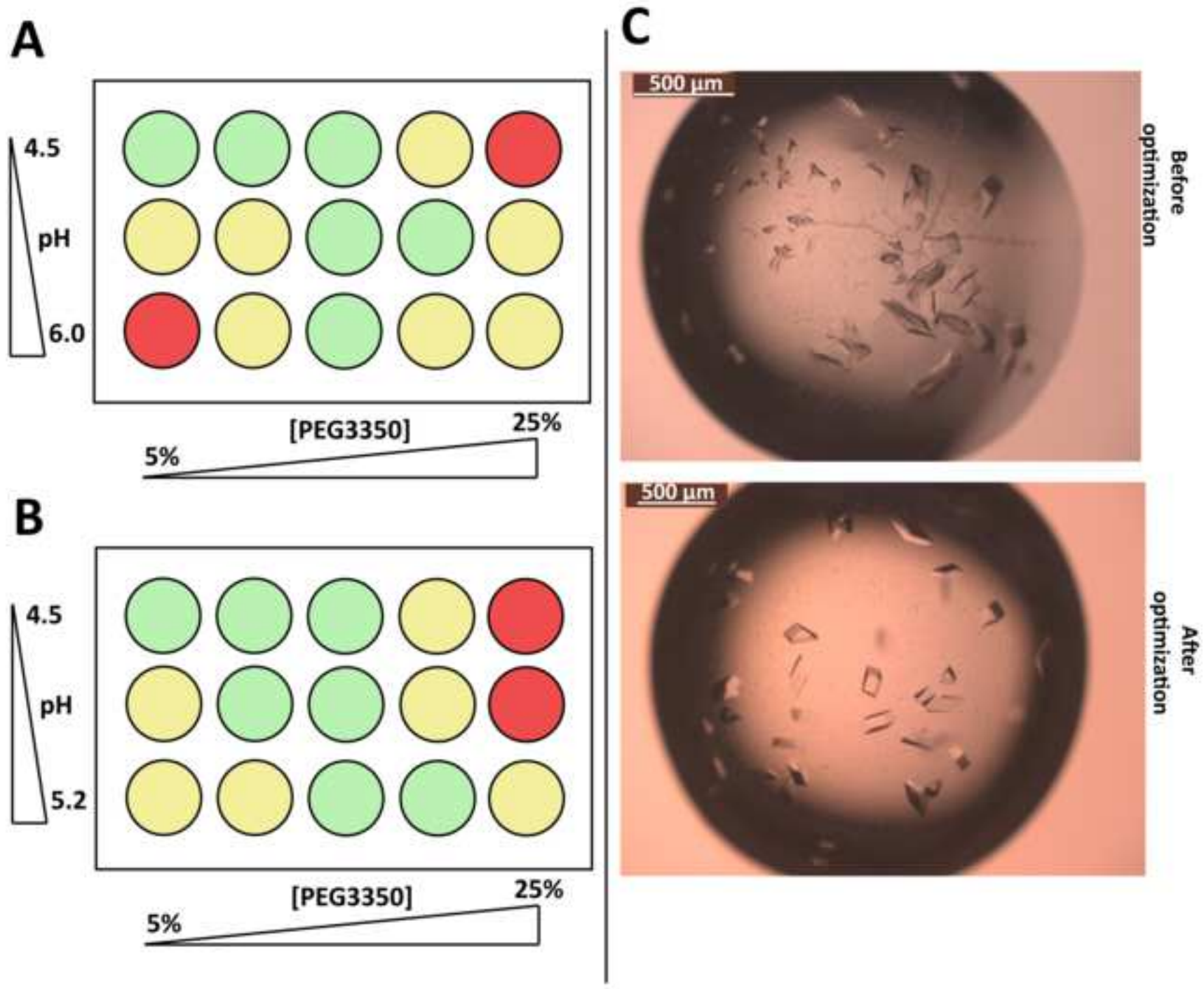

Figure 5

Click here to access/download;Figure;Figure5.psd

	LATTICE-CHARACTER	BRAVAIS-LATTICE	QUALITY OF FIT	UNIT CELL CONSTANTS (ANGSTROEM & DEGREES) a	b	c	alpha	beta	gamma	REINDEXING TRANSFORMATION											
*	31	aP	0.0	43.2	61.1	137.8	90.0	90.0	69.3	-1	0	0	0	0	1	0	0	0	0	-1	0
*	44	aP	0.7	43.2	61.1	137.8	90.0	90.0	110.7	1	0	0	0	-1	-1	0	0	0	0	-1	0
*	29	mC	1.0	43.2	114.4	137.8	90.0	90.0	90.0	-1	0	0	0	-1	-2	0	0	0	0	1	0
*	34	mP	1.7	43.2	137.8	61.1	90.0	110.7	90.0	-1	0	0	0	0	0	1	0	1	1	0	0
*	39	mC	1.8	114.4	43.2	137.8	90.0	90.0	90.0	1	2	0	0	-1	0	0	0	0	0	1	0
*	38	oC	2.1	43.2	114.4	137.8	90.0	90.0	90.0	-1	0	0	0	-1	-2	0	0	0	0	1	0
	35	mP	251.1	61.1	43.2	137.8	90.0	90.0	110.7	0	-1	0	0	-1	0	0	0	0	0	-1	0

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR COMPARED expected	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
7.20	6918	566	569	99.5%	15.4%	15.6%	6914	16.27	16.1%	98.8*	-61	0.389	382
5.12	11789	951	951	100.0%	17.2%	16.1%	11789	15.45	18.0%	98.5*	-45	0.524	754
4.19	13849	1200	1200	100.0%	18.2%	16.0%	13849	14.95	19.0%	97.9*	-30	0.639	1002
3.63	17555	1386	1386	100.0%	20.2%	17.2%	17555	14.00	21.1%	98.2*	-28	0.693	1201
3.25	21012	1570	1570	100.0%	21.7%	19.6%	21012	11.98	22.6%	98.6*	-26	0.625	1369
2.97	23980	1718	1718	100.0%	23.2%	25.5%	23980	9.24	24.1%	98.3*	-27	0.531	1518
2.75	25165	1854	1854	100.0%	28.8%	38.2%	25165	6.27	29.9%	97.0*	-19	0.483	1666
2.57	26007	1999	1999	100.0%	40.3%	55.8%	26007	4.16	41.9%	97.4*	-22	0.456	1799
2.43	25612	2052	2061	98.6%	48.9%	93.3%	25596	2.62	50.9%	92.4*	-29	0.374	1829
total	171885	13296	13329	99.8%	19.5%	19.2%	171867	8.12	20.3%	99.0*	-29	0.520	11520

NUMBER OF REFLECTIONS IN SELECTED SUBSET OF IMAGES	176795
NUMBER OF REJECTED MISFITS	4812
NUMBER OF SYSTEMATIC ABSENT REFLECTIONS	98
NUMBER OF ACCEPTED OBSERVATIONS	171885
NUMBER OF UNIQUE ACCEPTED REFLECTIONS	13296

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR COMPARED expected	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
7.04	7597	1058	1070	98.9%	9.4%	10.6%	7595	10.00	10.1%	99.4*	-49	0.409	981
5.01	12628	1860	1860	99.6%	11.1%	11.0%	12628	16.21	12.0%	98.4*	-32	0.584	1725
4.10	14550	2346	2364	99.2%	12.0%	10.9%	14550	15.38	13.1%	97.7*	-21	0.736	2104
3.55	18699	2800	2817	99.7%	13.9%	12.1%	18699	13.03	15.1%	97.9*	-24	0.757	2603
3.18	22304	3181	3190	99.7%	15.7%	14.5%	22303	11.38	17.0%	98.3*	-20	0.710	3044
2.91	25451	3477	3482	99.9%	19.0%	20.2%	25450	8.34	20.5%	98.0*	-19	0.623	3379
2.69	26408	3821	3831	99.7%	25.7%	32.2%	26408	5.28	27.8%	96.0*	-16	0.568	3663
2.52	27171	4086	4091	99.9%	36.7%	45.6%	27171	3.39	38.9%	92.2*	-14	0.563	3896
2.37	28434	4265	4358	97.9%	51.3%	82.8%	28359	2.12	55.6%	84.1*	-20	0.480	3967
total	183322	26902	27071	99.4%	13.6%	14.0%	183243	8.64	14.7%	99.2*	-21	0.606	25362

NUMBER OF REFLECTIONS IN SELECTED SUBSET OF IMAGES	189041
NUMBER OF REJECTED MISFITS	5617
NUMBER OF SYSTEMATIC ABSENT REFLECTIONS	102
NUMBER OF ACCEPTED OBSERVATIONS	183322
NUMBER OF UNIQUE ACCEPTED REFLECTIONS	26902

Figure 6

Click here to access/download;Figure;figure6.psd

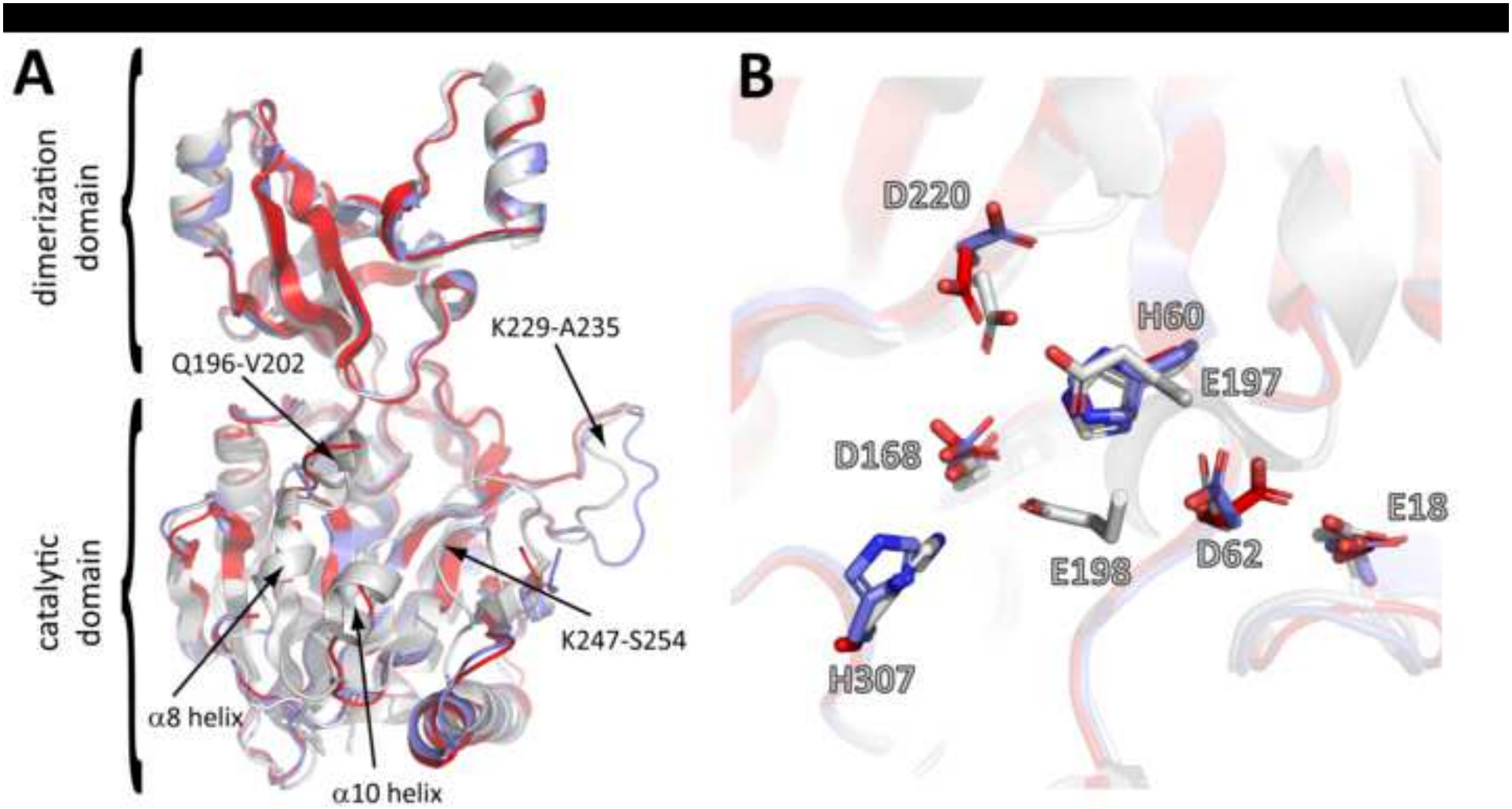

Table 1

	TmPep1050$_{H60A\ H307A}$
Data collection	
Temperature (K)	100
Radiation source	SOLEIL Proxima 2
Wavelength (Å)	0.9801
Detector	Dectris Eiger X 9M
Oscillation range (°)	0.1
Exposure time (s)	0.025
Space group	P 1 21 1
Unit cell parameters	
α, β, γ (°)	90.00, 110.69, 90.00
a, b, c (Å)	43.24, 137.79, 61.11
Resolution	43.99 – 2.37 (2.52-2.37)
Unique reflections	26.902
R_{merge} (%)	0.14
Redundancy	6,815
<I/σ>	8.64 (2.12)
Completeness (%)	99.6 (97.9)
$CC_{1/2}$ (%)	99.2 (84.1)
Refinement	
Resolution	43.99 – 2.37
Reflections	26.9
R_{free} set test count	1345
R_{work}/R_{free}	0.206/0.234
Protein molecules per ASU	2
V_M (Å^3/Da)	2.37
Solvent content (%)	49.0
Protein/solvent atoms	4,559/96
r.m.s.d. bond lengths (Å)	0.31
r.m.s.d. bond angles (°)	0.51
Average B-factors (Å^2)	57.0
Favored/disallowed Ramachandran φ/ψ (%)	95.02 / 0.17
Twin law	h, -k, -h-l
PDB code	5NE9

Name of Material/Equipment	Company	Catalog Number	Comments/Description
1,10-phenanthroline	Sigma-Aldrich	13, 137-7	
Amicon Ultra 0.5 ml Centrifugal Filters Ultracel 30K	Merck Millipore	UFC503096	
Amicon Ultra 15 Centrifugal Filters Ultracel 30K	Merck Millipore	UFC903024	
Benzonase Nuclease	Merck Millipore	70664-3	
CCP4	N/A		visit http://www.ccp4.ac.uk/
cOmplete EDTA-free	Roche	5056489001	
Coot	N/A		visit https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/c
Crystal Screen I	Hampton Research	HR2-110	
Crystal Screen II	Hampton Research	HR2-112	
DreamTaq Green PCR Master Mix	ThermoFisher Scientific	K1082	
EasyXtal 15-well tool	NeXtal	132007	
Escherichia coli PPY strain	N/A		see reference 31
Escherichia coli XL1 blue strain	Agilent	200249	
Gel Filtration Calibration Kit HMW	GE Healthcare Life Sciences	28-4038-42	
Gel Filtration Calibration Kit LMW	GE Healthcare Life Sciences	28-4038-41	
Gel Filtration Standard	Biorad	1511901	
GeneJET Plasmid Miniprep Kit	ThermoFisher Scientific	K0503	
Index	Hampton Research	HR2-144	
Litholoops	Molecular Dimensions		
L-leucine-p-nitroanilide	Bachem AG	40010720025	
Natrix 1	Hampton Research	HR2-116	

Natrix 2	Hampton Research	HR2-117	
Neggia plugin	Dectris	N/A	visit https://www.dectris.com/
NeXtal Tubes JCSG Core Suite I	NeXtal	130724	
NeXtal Tubes JCSG Core Suite II	NeXtal	130725	
NeXtal Tubes JCSG Core Suite III	NeXtal	130726	
NeXtal Tubes JCSG Core Suite IV	NeXtal	130727	
pBAD-TOPO	ThermoFisher Scientific	K430001	
Phenix	N/A		visit https://www.phenix-online.org/
Phusion High-Fidelity DNA polymer	ThermoFisher Scientific	F-530L	
Salt RX 1	Hampton Research	HR2-107	
Salt RX 2	Hampton Research	HR2-109	
SnakeSkin Dialysis Tubing, 3.5K MWCO	ThermoFisher Scientific	88242	
Source 15Phe	GE Healthcare Life Sciences	17014702	
Source 15Q	GE Healthcare Life Sciences	17094705	
Superdex 200 prep grade	GE Healthcare Life Sciences	17104301	
Thermotoga maritima MSB8 strain	American Type Culture Collection	ATCC 43589	
TmCD00089984	DNASU Plasmid Repository	N/A	
XDS	N/A		visit http://xds.mpimf-heidelberg.mpg.de/
xdsme	N/A		visit https://github.com/legrandp/xdsme

:oot/

Dear Xiaoyan Cao,

We are pleased to resubmit our manuscript entitled "" in JoVE. The editorial and reviewers' comments have been helpful for improving the manuscript. We've taken them into consideration to the best of our ability. We hope that the current version will meet JoVE standards and the reviewers' expectations. You'll find hereafter point-by-point answers to the comments.

Some parts of the manuscript have been substantially reorganized like the crystallization and X-ray diffraction sections. Three figures (Fig. 2 - 4) have been adapted in accordance to the comments. The highlighted steps have been revised to gain coherence for filming.

Thank you for considering our manuscript.

Yours sincerely,

Raphaël Dutoit

Point-by-point answers to editorial comments

Comment 1: Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.

We took this opportunity to proofread and correct several errors.

Comment 2: JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials. You may use the generic term followed by "(Table of Materials)" to draw the readers' attention to specific commercial names. Examples of commercial sounding language in your manuscript are: Phusion, ThermoFisher Scientific, DreamTaq, MerckMillipore, cOmplete, Roche, Source 15Q resin (GE Healthcare), etc.

We appreciate that JoVE applies such a guideline to make articles "brand-free". So, we deleted as mush as possible the references to commercial names. We have maintained only two references in the legends of Figure 1 and Figure 3. In Figure 1, the plasmid name is mentioned (pBAD). In Figure 3, the resin name (Superdex200) is mentioned because the observed elution volumes and the calibration curve greatly depend on the chromatography support.

Comment 3: 1.1.2: Please specify PCR conditions as well.

PCR conditions were added.

Comment 4: 1.1.6: Please specify the voltage used to run the agarose gel.

Electrophoresis voltage was added.

Comment 5: 3.1.4: Please provide the composition of crystallization solution. Is it included in the commercial kit?

The composition of crystallization solutions used for screening is included in the commercial kit. We added the mention in the revised manuscript to refer to manufacturer's guide and score sheet.

Comment 6: 3.2.1-3.2.4, 4.1.2, etc.: Please do not generalize the protocol and be as specific as you can with respect to your experiment providing all the details.

Indeed, we generalized the protocol for the crystallization and X-ray data acquisition as these steps must be adapted according to the protein of interest. To answer to this comment, we refocused the crystallization and data acquisition parts on what we did with TmPep1050, using TmPep1050$_{H60A\ H307A}$ as an example. We hope that the protocol is less general.

Comment 7: 4.2: The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections. Please describe in the imperative tense how to collect data and ensure that individual steps contain only 2-3 actions per step and a maximum of 4 sentences per step. Please move the discussion about the protocol to the Discussion.

We changed the 4.2. section in discrete steps.

Comment 8: Please ensure that the highlighted steps form a cohesive narrative with a logical flow from one highlighted step to the next. The highlighted text must include at least one action that is written in the imperative voice per step. Notes cannot usually be filmed and should be excluded from the highlighting.

We reviewed the highlighted steps to improve narrative coherence. We would have liked to include data acquisition in filming but, unfortunately, we have the localization constraint, this experimental part being done at the SOLEIL synchrotron in Paris.

Comment 9: Please include all relevant details that are required to perform the step in the highlighting. For example: If step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be highlighted.

We highlighted the sub-steps as well as.

Comment 10: When reviewing the highlighting length for the protocol, please watch out for repeated steps. Please ensure that the repeated step has been highlighted previously.

We highlighted other sections to meet this requirement. For instance, section 2.2 refers to an enzymatic assay which has included in the highlighted text (section 2.1.).

Comment 11: Figure 2 and Figure 3: Please abbreviate liters to L (mL) to avoid confusion.

Figure 2 and Figure 3 were adapted as required.

Comment 12: Figure 4: Please define the error bars in the figure legend. The numbers in the figure can be hardly read.

Indeed, the scale is almost invisible and was adapted (bigger line width and font size).

Point-by-point answers to Reviewer #1

Comment 1: In line 195, as wild-type TmPep1050 is dodecameric in solution, why not use concentrators with higher cut-off size here?

Indeed, we could have used an ultrafiltration device with 100-kDa cut-off. However, we decided to standardize the purification procedure so that it can be applied to any TmPep1050 variants. To avoid missing low molecular weight oligomers, we chose ultrafiltration units with 30-kDa cut-off.

Comment 2: In figure 3, it was highly recommended to add a gel filtration standard curve using the same column as molecular weight markers.

We totally agree with Reviewer 1's pertinent critic. We added a standard curve for gel filtration in Figure 3 as a new panel (D). The markers were analyzed in the same conditions as those of TmPep1050 samples.

Comment 3: In crystallization method, the temperature was absent. This parameter is critical and should be included in revised manuscript.

We added the temperature in the crystallization section.

Comment 4: In line 286, authors mentioned that crystals might appear in several minutes to months. However, it was of more interest when the crystals would reach their full size.

We were a bit inexplicit for describing the crystallization as the first version of our manuscript. This part has been rewritten to meet editorial comments. We added the notion of time for crystal appearance and growth.

Comment 5: There are a few ambiguous descriptions of concentration using percentage, for example, in line 153 "NaCl 0.9%" and in line 292 "40% PEG 3350". I guess they were both prepared in weight/volume manner, right? However, authors should point it out here.

Indeed, these solutions were prepared in weight/volume. We clarified it in the revised manuscript.