# Why are the 2-oxoacid dehydrogenase complexes so large? Generation of an active trimeric complex.

Nia L. Marrott[1], Jacqueline J.T. Marshall[2], Dmitri I. Svergun[3], Susan J. Crennell[1], David W. Hough[1], Jean M.H. van den Elsen[4]* and Michael J. Danson[1]*

[1] Centre for Extremophile Research, Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, U.K.
[2] Protein Phosphorylation Laboratory, London Research Institute, 44 Lincoln's Inn Fields, London, WC2A 3LY, U.K.
[3] EMBL, Hamburg Outstation, Notkestrasse 85, Hamburg D-22603, Germany.
[4] Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, U.K.

**Running title:** A trimeric 2-oxoacid dehydrogenase complex**.**

**Keywords:** Thermophile; multienzyme complex; macromolecular assembly; X-ray crystallography, Archaea

**Abbreviations:**
E1, 2-oxoacid decarboxylase; E2, dihydrolipoyl acyl-transferase; E2cat, C-terminal catalytic domain of the E2 component; E3, dihydrolipoamide dehydrogenase; $E_M$, molar absorption coefficient; OADHC, 2-oxoacid dehydrogenase multienzyme complex; PDB, Protein Data Bank; PDHC, pyruvate dehydrogenase complex; OGDHC, 2-oxoglutarate dehydrogenase complex; $R_{work}$, random errors of the working data set; $R_{free}$,random errors computed from a test set of reflections that was not used in the refinement; $R_g$, radius of gyration; SAXS, small-angle X-ray scattering; t-E2, truncated E2; t-E2cat, truncated E2cat

*Corresponding authors:

Professor Michael Danson
Centre for Extremophile Research
Department of Biology and Biochemistry
University of Bath
Bath
BA2 7AY, U.K.
Tel: +44-1225-386509
Fax: +44-1225-386779
Email: M.J.Danson@bath.ac.uk

Dr Jean van den Elsen
Department of Biology and Biochemistry
University of Bath
Bath
BA2 7AY, U.K.
Tel: +44-1225-383639
Fax: +44-1225-386779
Email: J.M.H.V.Elsen@bath.ac.uk

**Abstract**

The four component polypeptides of the 2-oxoacid dehydrogenase complex from the thermophilic archaeon, *Thermoplasma acidophilum*, assemble to give an active multienzyme complex possessing activity with the branched-chain 2-oxoacids derived from leucine, isoleucine and valine, and with pyruvate. The E2 core of the complex is composed of identical trimer-forming units that assemble into a novel 42-mer structure comprising octahedral and icosahedral geometric aspects. From our previously-determined structure of this catalytic core, the inter-trimer interactions involve a tyrosine residue near the C-terminus secured in a hydrophobic pocket of an adjacent trimer like a ball-and-socket joint. In the current paper, we have deleted the 5 terminal amino acids of the E2 polypeptide (IIYEI) and shown by equilibrium centrifugation that it now only assembles into a trimeric enzyme. This was confirmed by small-angle X-ray scattering analysis, although this technique showed the presence of approximately 20% hexamers. The crystal structure of the trimeric truncated E2 core has been determined and shown to be virtually identical to the ones observed in the 42-mer, demonstrating that removal of the C-terminal anchor does not significantly affect the individual monomer or trimer structures. The truncated E2 is still able to bind both E1 and E3 components to give an active complex with catalytic activity similar to the native multienzyme complex. This is the first report of an active mini-complex for this enzyme, and raises the question of why all 2-oxoacid dehydrogenase complexes assemble into such large structures.

## INTRODUCTION

The 2-oxoacid dehydrogenase multienzyme complexes (OADHCs) function in the pathways of central metabolism by catalysing the conversion of various 2-oxoacids, principally pyruvate, 2-oxoglutarate and the branched-chain 2-oxoacids, to their corresponding acyl-CoAs. These complexes all comprise multiple copies of three enzymes that catalyse consecutive reactions in the overall oxidative decarboxylation of their 2-oxoacid substrates: 2-oxoacid decarboxylase (E1), dihydrolipoamide acyltransferase (E2) and dihydrolipoamide dehydrogenase (E3) [1-3].

The E2 protein forms the structural core of the OADHC, and itself is composed of three structural domains separated by flexible linkers regions. The N-terminal lipoyl domain, with its covalently-attached lipoic acid moiety, is responsible for the transfer of the reactive intermediates between the active sites of E1, E2, and E3. Next to this in the sequence is the peripheral subunit-binding domain to which the E1 and E3 enzymes associate non-covalently; it is a small domain consisting of two short α-helices and a $3_{10}$-helix. The C-terminal catalytic domain of E2 is the largest of the three, and contains the acyltransferase active site and is also solely responsible for forming the structural core of the complex. To form this core, the E2 enzyme is thought to associate into trimers, and then 8 or 20 of these trimers associate into an octahedral (24-mer) or icosahedral (60-mer) core, respectively, with the flexible linkers and the lipoyl and subunit-binding domains projecting outwards. The oligomeric nature of the core assembly is dependent on both the complex and the source organism [4].

The genes encoding a 2-oxoacid dehydrogenase complex have been identified in the thermophilic archaeon *Thermoplasma acidophilum*. When recombinantly expressed, the resulting proteins assemble into a functional complex that has optimal activity with the branched-chain 2-oxoacids and, to a lesser extent, with pyruvate [5]. The structure of the E2 catalytic core has been determined to be a novel 42-mer structure [6]. The 24-mer and 60-mer E2 structures comprise only square or pentagonal faces, respectively, whereas the 42-mer has both within the one structure. Comparisons of the trimers from these various E2 assemblies show that the variation in structure is a result of how the E2 trimers interact with one another rather than changes in the structure of the trimers. It would therefore appear that the *T. acidophilum* E2s are capable of interacting differently with one another within a single assembled structure.

From the structures of the octahedral and icosahedral cores of OADHCs, it has been suggested that the C-terminal $3_{10}$-helix of the E2 is important in the formation of trimer-trimer interactions [4,7,8]. One residue in particular, an "anchor residue", is thought not only to be important in the formation of trimer-trimer interactions, but also to affect the angle at which the trimers interact, and this in turn influences the multimeric structure of the E2 assembly [4]. This E2 anchor residue is a hydrophobic amino acid, usually a valine or leucine, located a few amino acids from the E2 C-terminus; the side-chain of the hydrophobic anchor residue extends from one trimer into a hydrophobic pocket on the adjacent trimer, and *vice versa,* tethering neighbouring trimers to each other. Removal of the five C-terminal residues from porcine E2 of the 2-oxoglutarate complex (OGDHC) prevented cubic assembly but also abolished all E2 activity [9]. Moreover, although a trimeric E2 from the *Geobacillus stearothermophilus* pyruvate dehydrogenase complex (PDHC) has been reported [10], the effects on E2 and OADHC activities have not, to date, been published.

From a comparison of the *T. acidophilum* E2 structure with the E2 structures from the PDHCs of *G. stearothermophilus, Enterococcus faecalis* and *Azotobacter vinelandii*, residue Y398 within the five C-terminal amino acids of the *T. acidophilum* enzyme (396-400: IIYEI) is secured in a hydrophobic pocket of the opposing trimer, like a ball-and-socket joint [6] (Figure 1).

In the current paper, through the introduction of a stop codon into the gene encoding the *T. acidophilum* E2 protein, we have generated an E2 lacking the five C-terminal amino acids, hereafter referred to as a truncated E2 (t-E2). The resulting enzyme is a trimer that is catalytically active when recombinantly expressed, and is capable of binding both E1 and E3 proteins to produce a complex that catalyses the oxidative decarboxylation of branched-

chain 2-oxoacids. The crystal structure of the trimeric catalytic truncated E2 core (t-E2cat) is also reported.

This trimeric *T. acidophilum* OADHC is the first reported functional mini-complex, and poses the question of why Nature has evolved such large multienzyme complexes when smaller assemblies are capable of catalysing the same 2-oxoacid decarboxylations.

## METHODS

### Removal of the E2 C-terminus by insertion of a premature stop codon

Using the Stratagene Quick-change site-directed mutagenesis protocol, the five C-terminal amino acids of the *T. acidophilum* E2 were removed by the insertion of a premature stop codon (TGA) into the previously-cloned gene encoding whole E2 in the pET28a vector [5]. The same mutagenesis was carried out on the gene fragment encoding just the E2 C-terminal catalytic domain (E2cat) in pET24a [6]. DNA sequencing confirmed the presence of the mutation in both cases.

Both truncated E2 proteins (t-E2 His-tagged and t-E2cat with no tag) were expressed in *E. coli* Rosetta (DE3) grown at 37°C in LB supplemented with kanamycin (30 µg/ml) and chloramphenicol (34µg/ml), and induced with 1 mM IPTG. The medium was supplemented with 0.2 mM DL-lipoic acid when lipoylated E2 was required. Cells were harvested 20 h after induction and lysed by sonication. The t-E2 was purified by nickel affinity chromatography, with proteins eluted by an imidazole gradient ranging from 20 mM to 1 M; the t-E2cat was purified by anion exchange on HiTrap Q (HP), followed by gel filtration on a Superdex 200 column, as described in [6].

### Expression of the recombinant *T. acidophilum* E1 and E3 enzymes

The previously-cloned E1α (pET19b) and E1β (pET24a) genes were co-expressed in *E. coli* Rosetta (DE3), and the recombinant $\alpha_2\beta_2$ enzyme was purified as described in [5]. Recombinant E3($\alpha_2$) was produced as described in [5].

### Assay of E2 catalytic activity

The dihydrolipoamide S-acetyltransferase activity of the t-E2 and t-E2cat enzymes was determined using a coupled assay with *Geobacillus stearothermophilus* phosphotransacetylase [11]. The 1 ml assay contained 0.1 M Tris-HCl (pH 7.6 at 20°C), 4.0 mM DL-dihydrolipoamide, 0.1 mM Coenzyme-A (CoA), 10 mM acetyl-phosphate and 7.5 U phosphotransacetylase; the phosphotransacetylase reaction producing acetyl-CoA was allowed to reach equilibrium at 55°C before the E2 enzyme was added. The increase in thioester bond concentration due to the E2-catalysed formation of acetyl-lipoamide at 55°C was followed at 233 nm ($E_M$ = 5.4 x $10^3$ $M^{-1}$ $cm^{-1}$).

### Assembly and assay of OADHC containing the truncated E2

The lipoylated t-E2 was used to assemble an OADHC. The component enzymes E1$\alpha_2\beta_2$, t-E2, and E3$\alpha_2$ were mixed in the required molar stoichiometry, diluted 1:1 with 50 mM potassium phosphate pH 7.0, 2.5 mM $NAD^+$, 1 mM $MgCl_2$, and 0.2 mM TPP, and the mixture was incubated at 55°C. A 1 ml assay contained 50 mM potassium phosphate pH 7.0, 2.5 mM $NAD^+$, 1 mM $MgCl_2$, 0.2 mM TPP, 0.13 mM CoA, 2.6 mM cysteine/HCl, and 2 mM 3-methyl-2-oxopentanoic acid. The rate of NADH production was measured spectrophotometrically at 340 nm ($E_M$ = 6.22 x $10^3$ $M^{-1}$ $cm^{-1}$).

### Analytical equilibrium centrifugation

Three different concentrations of the t-E2 and t-E2cat in 50 mM Tris-HCl, pH 8.8, 100 mM NaCl, plus buffer-only references, were loaded into 6-sector centre-pieces within centrifuge cells. Protein monomer concentrations of t-E2 were 28, 18 and 8.6 µM (1.3, 0.81 and 0.36 mg/ml), and for t-E2cat were 62, 24 and 10 µM (1.5, 0.6 and 0.25 mg/ml), as calculated from their absorbance values recorded immediately upon starting centrifugation at 3,000 rpm

(pathlength 1.2 cm) using $E_M$ values of 23,380 $M^{-1}.cm^{-1}$ (t-E2) and of 13,410 $M^{-1}.cm^{-1}$ (t-E2cat) [12]. Centrifugation was carried out in an An60Ti rotor in a Beckman XL-A ultracentrifuge at 20°C. At time intervals, the $A_{280}$ of the samples was measured at radial positions from 5.8 to 7.2 cm in 0.03 cm steps. Cells were spun at 7,000 and 9,500 rpm (t-E2) and at 12,000, 15,000 and 18,000 (t-E2cat) for 24-34 h each (longer times for lower speeds), with scans completed at 4-6 h intervals after an initial 18 h. Equilibrium was reached by each final scan time as confirmed by overlay of scans completed at the final and penultimate scan times. A final 6-hour clearing spin at 40,000 rpm, during which time the protein was removed from the meniscus of the solution, enabled the background absorbance of the buffer in the samples versus the reference to be measured.

At each speed, the data from the final scan were simultaneously fitted in ORIGIN (Beckman) to the model for a single ideal species using partial specific volumes of 0.741 ml/g (t-E2) and 0.748 ml/g (t-E2cat) and a buffer density of 1.004 g/ml (estimated from the protein sequence and the buffer composition, respectively, using SEDNTERP [http://www.jphilo.mailway.com/download.htm]). Offset values for each data-set were floated in this global fit, but they closely matched the absorbance values at the menisci obtained from the clearing spin. Error margins were obtained from a weighted fit to the same data and are from the 95% confidence limits from this fitting routine. The differences between the data-points in each data-set and the global fit to the six (t-E2) or nine (t-E2cat) data-sets, the residuals, were also calculated.

**Small-angle X-ray scattering analysis**

Synchrotron radiation X-ray scattering data collection was performed at the X33 beam line of the EMBL Hamburg Outstation (DORIS III storage ring at DESY [13]), as described previously [6]. Solutions of t-E2cat were measured at protein concentrations of 3.2, 4.3 and 5.8 mg/ml at the sample temperature 10˚C on a Pilatus-1M detector at the X-ray wavelength λ = 1.5 Å, covering the momentum transfer range $0.1 < s < 0.6$ $nm^{-1}$ ($s = 4\pi \sin(\theta)/\lambda$ where $2\theta$ is the scattering angle). Prior to data collection, the monodispersity of the protein solutions was ensured using dynamic light scattering (Nano-SZetasizer, Malvern, UK). The data were reduced and processed by PRIMUS [14] and GNOM [15] as described in [6] to obtain a composite scattering pattern by appropriately merging the higher and lower concentration data and also to compute the overall particle parameters, radius of gyration, $R_g$, maximum size $D_{max}$ and the excluded volume $V_p$. The scattering patterns from the high resolution models were computed with CRYSOL [16] minimising discrepancy to the experimental data $I_{exp}(s)$:

$$\chi^2 = \frac{1}{N-1}\sum_j \left[\frac{I_{\exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2$$

where N is the number of experimental points, c is a scaling factor, and $I_{calc}(s)$ and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer $s_j$, respectively. The rigid body modelling, using the crystal structure of the truncated E2cat trimer and accounting for possible presence of hexameric species, was done using the program SASREFMX [17]. In an alternative approach, available models of hexameric protein were used in the program OLIGOMER [14] to calculate the volume fractions of monomers and hexamers in the putative mixtures.

**Crystallisation, data collection and structural analysis**

Crystals of truncated E2cat were grown in conditions containing 8% (w/v) PEG 4k, 3% (w/v) PGA, and 0.1 M Na cacodylate, pH 6.5, using the 'hanging drop' vapour diffusion method. X-ray diffraction data were collected at the Diamond Light Source (Oxfordshire, UK) with 30% (w/v) glycerol as cryoprotectant, and were processed using the HKL2000 package. Molecular replacement was carried out with BALBES [18], using a trimer from the structure of the non-truncated *T. acidophilum* E2cat (Protein Data Bank accession code 3RQC) as a search model. Model building was done with COOT [19] followed by rounds of refinement

using REFMAC5, part of the CCP4 software package [20]. The final $R_{work}$ and $R_{free}$ values were 21.5% and 28.4%. The final coordinates of the t-E2cat structure have been deposited to the Protein Data Bank (PDB accession code: 4OFS). The E2cat and t-E2cat crystal structures were compared using the DALILITE programme [21].

## RESULTS

### Characterisation of the truncated E2 and E2cat proteins

The truncated versions of the His-tagged whole E2 and a non-tagged E2cat (lacking the peripheral subunit-binding domain and the lipoyl domain) were successfully expressed in *E. coli* and purified to homogeneity. Both enzymes were active with respect to their transacetylase activities, for which a lipoyl domain is not required, although the t-E2cat had lost a significant amount of its catalytic ability. Determined $k_{cat}$ values with respect to their non-truncated versions are shown in Table 1. Similar changes in the thermostabilities of the two proteins were observed; whereas the thermostability of the t-E2 was similar to that of the whole E2, the t-E2cat was far less stable than E2cat, which showed no loss of activity after 5 h.

Analytical equilibrium ultracentrifugation was used to determine the oligomeric assembly of the truncated E2 proteins. Fitting the data to a model of a single ideal species gave $M_r$ values recorded in Table 1; similar results were obtained when fitting data at each separate concentration alone. In both cases, removal of the five C-terminal amino acids of the E2 protein has resulted in a species not able to assemble into the 42-mer of the native E2 and E2cat enzymes, but to remain in solution as a trimeric species, consistent with the predicted basic building blocks of all OADHCs E2 cores.

### Crystal structure of the truncated E2cat trimer

The crystals of the *T. acidophilum* t-E2cat belong to trigonal space-group $P3_221$ and a complete data set was processed to 4.1 Å resolution. For X-ray data collection and model refinement statistics, see Tables S1 and S2. For structural determination, a molecular replacement solution was obtained using a trimer from the *T. acidophilum* E2cat structure (PDB accession code 3RQC), using the program BALBES [18]. The asymmetric unit of the t-E2cat contains 6 monomers arranged as a crystallographic dimer of trimers, which are very similar to each other (RMSD: 0.5 Å over 214/219 residues) and virtually identical to the ones observed in the E2cat 42-mer (RMSD: 0.8 Å (222/222), Figure 2), demonstrating that removal of the C-terminal anchoring helix does not significantly affect the individual monomer or trimer structures. The relative orientation of the two trimers in the asymmetric unit, however, is very different from the non-truncated E2cat trimers in the 42mer (Figure 3).

In the t-E2cat trimer one of the trimers has rotated almost 180˚ compared to its position in the E2cat 42mer structure (along the axis connecting the centres of the two trimers), and in this orientation could not pack into the 42-mer. Whilst in the E2cat 42mer structure the trimers interact via intricate hydrophobic ball-and-socket joint interactions between C-terminal residues of each monomer in the trimers (as described above, and labelled in Figure 3) the corresponding C-termini of the t-E2cat trimers are still in close proximity (~18 Å compared to ~11 Å in the 42mer) but do not share any interaction surface, as expected. The dimer of trimers is only loosely held together by hydrogen bonds between the $\alpha$2 helix of chain C and the $\alpha$3 helix of chain E. The t-E2cat trimer dimer interface shows a significantly reduced interaction surface of ~440 $Å^2$, compared to those observed in the 42mer, with buried surface areas of ~680 $Å^2$ and ~775 $Å^2$ (contributed by each of the trimers at the pentamer-pentamer and the square-pentamer interfaces, respectively).

### Small-angle X-ray scattering

The $R_g$ values determined from the experimental SAXS data demonstrated a weak concentration dependence, decreasing with the solute concentration (Table 2). All the experimental $R_g$ values noticeably exceeded the value computed from the crystal structure of the t-E2cat trimer (2.6 nm). The $M_r$ of the solute, calculated from the forward scattering

and from the excluded particle volume, was 80±8 kDa and 85±8 kDa, respectively, which exceed the $M_r$ of the trimer (74.5 kDa). Furthermore, the scattering pattern computed from the crystal structure of t-E2cat trimer yielded a high discrepancy $\chi$ = 2.1 to the composite experimental curve by t-E2cat. As is evident from Figure 4A (insert), the scattering computed from the trimer displays major deviations from the experimental data at low angles, which reflect the overall structure. At higher angles, reflecting the internal structure, the computed curve provides a reasonable fit (Figure 4A). The above data therefore strongly suggest that t-E2cat does not form the 42-mer assembly observed for E2cat. Instead, the t-E2cat trimer is the dominant species in solution, but the sample most probably contains some proportion of higher oligomers of this protein.

To test this hypothesis further, the scattering data at different concentrations were fitted using a rigid body SASREFMX, which builds models of t-E2cat hexamer as a symmetric assembly of two trimers. The program allows for partial dissociation of the hexamers and also determines the volume fraction of the dissociated species in solution. Several SASREFMX runs consistently yielded t-E2cat hexamers as side-by-side associates of trimers. The volume fractions of the hexamers in the mixtures were around 20-25%, increasing with concentration (Table 2). Interestingly, the hexamer revealed by SASREFMX has the overall shape similar to the dimer of trimers observed in the t-E2cat crystal structure (Figure 4B). Allowing for the presence of hexamers significantly improved the fit to the experimental data and Figure 4A displays the fit computed from a mixture of 75 volume percent trimers and 25 volume percent crystallographic t-E2cat hexamers, yielding a discrepancy of $\chi$=1.2.

**Assembly and characterisation of a trimer OADHC**

As reported above, the trimeric t-E2 has retained approximately 90% of the transacetylase activity of the non-truncated E2 that is part of the native 42-mer structure. It has been shown in the E2 of the *A. vinelandii* pyruvate dehydrogenase complex that the transacetylase active site is located within a channel situated at the intra-trimer monomer faces [7,8]. This channel is present and unchanged in the t-E2. Moreover, as only the five C-terminal amino acids have been removed, the t-E2 retains its peripheral subunit binding domain and its lipoyl domain, and the possibility that E1, trimeric t-E2 and E3 might associate to form an active mini-complex was therefore investigated.

The t-E2 was recombinantly expressed in *E. coli* grown in a medium supplemented with 0.2 mM DL-lipoic acid, under which conditions we have previously shown that the whole E2 is lipoylated. The purified E2 was then mixed with purified E1 and E3 in a molar ratio of 1:1:0.5, as described in Material and Methods. The mixture was passed down a Sepharose gel-filtration column, and a protein species was collected in the high molecular size region ($M_r$ >800,000). SDS-PAGE analysis confirmed that it contained E1, E2 and E3 enzymes, demonstrating that the trimeric t-E2 was capable of binding the other component enzymes. Using the branched-chain 2-oxoacid, 3-methyl-2-oxopentaonic acid, which is a substrate for the native 42-mer complex, the 'trimeric' complex was shown to be catalytically active and its kinetic parameters were determined under identical conditions to the 42-mer complex formed by mixing components in the same stoichiometry (Table 3).

The binding of E1 to non-truncated (42-mer) and truncated (trimer) E2s appeared to be very similar, each reaching maximal complex activity at a mixing ratio of E1:E2 of 1.0-1.5 (Figure 5). Additionally, the temperature optima of the two complexes were similar (Figure 6), although the thermostabilities, as measured by the resistance to irreversible thermal inactivation, were different, with the 42-mer complex being significantly more stable than the trimer (Table 3).

**DISCUSSION**

The 42-mer E2 core of the *T. acidophilum* 2-oxoacid complex comprises identical trimer-forming units with a unique oblate spheroid geometry, comprising an amalgam of octahedral and icosahedral geometric aspects [6]. However, the trimer building blocks in the archaeal complex are similar to those in the cubic and dodecahedral E2 core structures of bacteria and eukaryotes, and the differences in the assembly are a result of how the E2 trimers interact with one another. None-the-less, it appears that there is some commonality to the trimeric interactions across all E2 structures, in that a hydrophobic anchor residue near the C-terminus extends from one trimer into a hydrophobic pocket on the adjacent trimer, and *vice versa,* tethering neighbouring trimers to each other. In the *T. acidophilum* complex the hydrophobic residue appears to be Y398 [6], three amino acids from the C-terminus (Figure 1), and therefore in the present paper we have deleted the last 5 residues from the E2 polypeptide and observed the effect on the core assembly.

The removal of the 'ball-and-socket' hydrophobic interaction does indeed prevent the assembly into a 42-mer, with equilibrium centrifugation indicating only a trimeric species whilst SAXS analysis finding a small amount of hexamers, albeit at a higher protein concentration than the highest used in the centrifugation. Interestingly, we have previously observed in the 42-mer structure an ionic interaction between R292 of the E2 polypeptide and its neighbouring residue D291 in the two-fold symmetry-related chain [6], and proposed that it may contribute to the pentameric interactions. However, at the pentamer-quadrilateral interfaces the distance between R292 and D291 is significantly larger than at the pentamer-pentamer interfaces, preventing an ionic interaction. The ionic interaction may account for some of the trimers remaining weakly associated in pairs to former hexamers, as observed by SAXS, whereas the lower protein concentrations used in the centrifugation experiments might result in complete dissociation into trimers.

Similar mutational effects on the assembly of E2 trimers have been reported for porcine E2 of OGDHC [9] and for the E2 of *G. stearothermophilus* PDHC [10]. However, only in the *T. acidophilum* situation reported here is the t-E2 catalytically active, as measured by its transacetylase activity [22]. Moreover, not only had the trimer retained this activity, it also bound the E1 and E3 components in a similar manner to the fully-assembled 42-mer, to give a 'mini' multienzyme complex with OADHC catalytic activity.

This retention of function by the trimeric E2 suggests no significant change in the structure of the enzyme, and this was confirmed by comparison of the crystal structures of the t-E2 catalytic trimer and the trimer within the 42-mer catalytic core (Figure 2). From the structure of the *A. vinelandii* PDHC E2 cubic core and site-directed mutagenesis experiments [7 and references therein], the active site residues of the E2 transacetylase are positioned within a tunnel between the monomers of each trimer. These residues in the *T. acidophilum* enzyme can be predicted based on sequence alignments, and three functional residues within the active site are found in the same positions in the t-E2 trimer and the 42-mer trimers: T327 from one monomer and H371 and D375 from an adjacent monomer (whole E2 numbering; equivalent to T151, H195 and D199 of t-Ecat [PDB accession code 4OFS]).

The question remains as to why all ODHCs assemble into such large structures when the data in this paper have demonstrated that a smaller trimeric complex can be functional. Increased stability may be one factor, as the trimeric complex is more thermolabile than the complex based on the 42-mer even though the two showed similar temperature optima. However, there are numerous examples of highly thermostable proteins that are monomers or small oligomers. A more plausible explanation lies in the phenomenon of active-site coupling that has been found in many OADHCs. This was first observed in the *E. coli* PDHC [23] and involves the transfer of acetyl-groups from one lipoyl residues to another around the complex core, such that the whole core can be acetylated from pyruvate entering via one E1 enzyme. In this way, pyruvate and CoASH can be linked irrespectively of where they bind to the complex. This coupling of active sites was shown to be rapid [24], providing a mechanism for rate enhancement in a multimeric structure when substrates (in that case, pyruvate and CoASH) are at low concentrations. The phenomenon

has now been demonstrated in a range of OADHs and provides the one rationale for a large assembly.

## Author Contributions

Michael Danson and David Hough planned and supervised the project and, with Nia Marrott, designed the experiments. Jacqueline Marshall carried out the ultracentrifugation analyses, and Dimitri Svergun carried out the SAXS analyses. All other laboratory experiments, including the mutagnesis, gene expression and the purification and characterisation of the recombinant proteins were conducted by Nia Marrott. Susan Crennell and Jean van den Elsen collected and processed the X-ray crystallographic data, calculated initial structures and guided Nia Marrott in the refinement of those structures. All the authors contributed to the writing of the paper.

## Acknowledgements

We are grateful to the UK Biotechnology & Biological Research Council for financial support in the form of a Studentship to NLM. Part of the work was supported by a Research Grant from the US Air Force Office of Scientific Research. The analytical ultracentrifugation data were collected with equipment provided by Professors S.E. Halford FRS and D.N. Woolfson, University of Bristol, U.K., and we thank them for help and valuable discussions. This work was carried out with the support of the Diamond Light Source, and access to the DESY-EMBL beam lines was supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 226716.

## References

1. Perham, R.N. (1991) Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes - a paradigm in the design of a multifunctional protein. Biochemistry USA **30**, 8501-8512
2. Perham, R.N. (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. Ann. Rev. Biochem. **69**, 961-1004
3. Perham, R.N., Jones, D.D., Chauhan, H.J. and Howard, M.J. (2002) Substrate channelling in 2-oxo acid dehydrogenase multienzyme complexes. Biochem. Soc. Trans. **30**, 47-51
4. Izard, T., Ævarsson, A., Allen, M.D., Westphal, A.H., Perham, R.N., de Kok, A. and Hol, W.G.J. (1999) Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. Proc. Natl. Acad. Sci. USA **96**, 1240-1245
5. Heath, C., Posner, M.G., Aass, H.C., Upadhyay, A., Scott, D.J., Hough, D.W. and Danson, M.J. (2007) The 2-oxoacid dehydrogenase multi-enzyme complex of the archaeon *Thermoplasma acidophilum* - recombinant expression, assembly and characterization. FEBS J. **274**, 5406-5415
6. Marrott, N.L., Marshall, J.J.T., Svergun, D.I., Crennell, S.J., Hough, D.W., Danson, M.J. and van den Elsen, J.M.H. (2012) The catalytic core of an archaeal 2-oxoacid dehydrogenase multienzyme complex is a 42-mer protein assembly. FEBS J. **279**, 713–723
7. Mattevi, A., Obmolova, G., Schulze, E., Kalk, K.H., Westphal, A.H., de Kok, A. and Hol, W.G. (1992) Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. Science **255**, 1544-1550
8. Mattevi, A., Obmolova, G., Kalk, K.H., Westphal, A.H., de Kok, A. and Hol, W.G. (1993) Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 26 Å resolution. J. Mol. Biol. **230**, 1183-1199

9. Koike, K., Suematsu, T. and Ehara, M. (2000) Cloning, overexpression and mutagenesis of cDNA encoding dihydrolipoamide succinyltransferase component of the porcine 2-oxoglutarate dehydrogenase complex. *Eur. J. Biochem.* **267,** 3005-3016
10. Peng, T., Lee, H. and Lim, S. (2012) Isolating a trimer intermediate in the self-assembly of E2 protein cage. Biomacromolecules **13**, 699-705
11. Packman, L.C., Perham, R.N. and Roberts, G.C.K. (1984) Domain structure and $^{1}$H-nmr spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. Biochem. J. **217**, 219-227
12. Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182**, 319-326
13. Blanchet, C.E., Zozulya, A.V., Kikhney, A.G., Franke, D., Konarev, P.V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., Svergun, D.I. and Roessle, M. (2012) Instrumental setup for high-throughput small- and wide-angle solution scattering at the X33 beamline of EMBL Hamburg. Journal of Applied Crystallography **45**, 489-495
14. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS - a Windows-PC based system for small-angle scattering data analysis. J. Appl. Crystallogr. **36**, 1277-1282
15. Svergun, D.I. (1992) Determination of the regularization parameter in indirect transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503
16. Svergun, D.I., Barberato, C. and Koch, M.H.J. (1995) CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. **28**, 768-773
17. Petoukhov, M.V., Billas, I.M., Takacs, M., Graewert, M.A., Moras, D. and Svergun, D.I. (2013) Reconstruction of quaternary structure from X-ray scattering by equilibrium mixtures of biological macromolecules. Biochemistry. **52**, 6844-55
18. Long, F., Vagin, A.A., Young, P. and Murshudov, G.N. (2008) BALBES: a molecular-replacement pipeline. Acta Crystallogr. D **64**, 125-132
19. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallog. **D60**, 2126-2132
20. Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. and Vagin, A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. **D67**, 355-367
21. Hasegawa, H. and Holm, L. (2009) Advances and pitfalls of protein structural alignment. Curr. Opin. Struct. Biol. **19**, 341-348
22. Packman, L.C., Perham, R.N. and Roberts, G.C.K. (1984) Domain structure and $^{1}$H-nmr spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus.* Biochem. J. **217**, 219-227
23. Bates, D.L., Danson, M.J., Hale, G., Hooper, E.A. and Perham, R.N. (1977) Self-assembly and catalytic activity of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli.* Nature **268**, 313-316
24. Danson, M.J., Fersht, A.R. and Perham, R.N. (1978) Rapid intramolecular coupling of active sites in the pyruvate dehydrogenase complex of *Escherichia coli*: mechanism for rate enhancement in a multimeric structure. Proc. Natl. Acad. Sci. USA **75**, 5386-5390

## Figure legends

**Figure 1. Detailed view of the hydrophobic inter-trimer interactions in the 42-mer E2 core assembly.**
In the crystal structure of the 42-mer E2 catalytic core [6], seven polypeptides (numbered A to G) are observed in the asymmetric unit. Six of these (A to F) form two tightly-packed trimers, and the figure shows the interactions between polypeptides C and E in the 42mer (at the pentamer-pentamer interface).

**Figure 2. Comparison of the trimeric structures of E2cat and the t-E2cat.**
The t-E2cat trimer (this work) is shown in cyan and this is overlaid with the non-truncated E2cat trimer, shown in black, taken from the 42-mer crystal structure [6].

**Figure 3. Comparison of the trimer-trimer interactions observed in the crystal structures of the E2cat 42mer (top) and the t-E2cat protein (bottom)**.
Shown are ribbon diagrams of the two trimers within the asymmetric unit. In both structures, the corresponding chains (chains A, B and C (left) and chains D, E and F (right)) within the trimers are rendered with identical colours for clarity. For the E2cat 42mer (top structure) the pentamer-pentamer interface is shown, and the C-termini involved in the trimer-trimer interaction are labelled (C).

**Figure 4. Tentative model of the *T. acidophilum* t-E2cat trimer-dimer observed in the SAXS analysis.**
(A) SAXS curves of the *T. acidophilum* t-E2cat in solution: (1) experimental data (composite curve, dots with error bars); (2) and (3) computed scattering from the crystallographic trimer and from its best mixture with the crystallographic hexamer as determined by OLIGOMER, respectively. (B) A tentative model of the *T. acidophilum* t-E2cat trimer-dimer observed in the SAXS analysis. The ribbon diagram of the t-E2cat trimer shows the position of the second trimer (left) constructed by fitting the SAXS data as a mixture of trimers. The two trimers are rendered using different types of ribbon for clarity. The color coding of the chains is identical to that used in Figure 3.

**Figure 5. Binding of E1 to non-truncated and truncated E2.**
Each E2 was mixed with E3 in a 1: 0.5 molar ratio, and these mixtures were then titrated with enzyme E1. Each mixture was assayed for whole complex activity using 3-methyl-2-oxopentaonic acid as the 2-oxoacid substrate. Complex activity is expressed as a percent value of the maximum activity achieved for each E2. (●) Non-truncated E2. (○) Truncated E2.

**Figure 6. Temperature optima of native and trimeric complexes.**
The assembled complexes (E1:E2:E3 stoichiometry of 1:1:0.5) were assayed over a range of temperatures using 3-methyl-2-oxopentaonic acid as the 2-oxoacid substrate. Activities are expressed as a percentage of the maximal activity. (●) Non-truncated E2. (○) Truncated E2.

**Table 1. Transacetylase activities, thermostability and assembly of the various E2s.**

| E2 variant | $k_{cat}$ ($min^{-1}$) | $t_{1/2}$ at 85°C (min) | Monomer $M_r$ (theoretical) | *Measured complex* $M_r$ | Number of monomers | References |
|---|---|---|---|---|---|---|
| E2 | 48.9 (± 2.3) | 40 | 46,408 | 1,940,000 (± 50,000) | 41.8 (± 1.1) | [6] |
| t-E2 | 37.8 (± 0.2) | 55 | 45,776 | 127,000 (± 3,000) | 2.8 (± 0.1) | This work |
| E2cat | 42.8 (± 3.4) | No loss of activity after 5 h | 25,455 | 1,080,000 (± 20,000) | 42.2 (± 0.8) | [6] |
| t-E2cat | 14.5 (± 2.0) | 125 | 24,823 | 70,200 (± 3,200) | 2.8 (± 0.1) | This work |

**Table 2 Concentration dependence of the trimer/hexamer ratio of truncated E2cat protein observed in the SAXS analysis**

| Protein concentration (mg/ml) | $\chi$ | Radius of gyration ($R_g$) (nm) | Trimer volume fraction ($V_{fr}$) | Hexamer volume fraction ($V_{fr}$) |
|---|---|---|---|---|
| 3.2 | 0.9 | 3.05± 0.05 | 0.78 ± 0.03 | 0.22± 0.03 |
| 4.3 | 1.1 | 3.13± 0.05 | 0.77 ± 0.03 | 0.23± 0.03 |
| 5.8 | 1.3 | 3.18± 0.05 | 0.74 ± 0.03 | 0.26± 0.03 |
| Composite data | 1.2 | 3.00± 0.05 | 0.75 ± 0.03 | 0.25± 0.03 |

**Table 3. Kinetic parameters and enzyme stability of the E2 variants in complex with E1 and E3.**

| E2 variant | Assembled structure | $K_m$ (mM) | $V_{max}$ (U/mg of E2) | $V_{max}/K_m$ (U/mg of E2/mM) | $t_{1/2}$ at 65°C (min) |
|---|---|---|---|---|---|
| Full-length E2 | 42-mer | 0.31 ± 0.01 | 0.92 ± 0.01 | 3.0 ± 0.1 | 290 |
| t-E2 | Trimer | 0.12 ± 0.01 | 0.43 ± 0.01 | 3.5 ± 0.1 | 40 |

The component enzymes were mixed in the cuvette in a stoichiometry of 1:1:0.5 ($E1\alpha_2\beta_2$:E2α:$E3\alpha_2$). The 2-oxoacid substrate was 3-methyl-2-oxopentaonic acid, and the specific activity is given as U/mg of E2, where 1 Unit is described as the conversion of 1 µmol $NAD^+$ to NADH per minute.

**Table S1 X-ray data collection statistics**

| | | |
|---|---|---|
| Space Group | | $P3_221$ |
| Cell parameters | a=b, c | 107.2 Å, 238.6 Å |
| | α=β, γ | 90°, 120° |
| Resolution (Å) (last shell) | | 4.1 (4.17-4.10) |
| Number of observations | | 12274 |
| Number of unique reflections | | 12902 |
| Completeness (last shell) | | 99.2% (97.4) |
| R merge (%)(last shell) | | 0.122 (0.551) |
| Average I/σI (last shell) | | 13.06 (2.48) |
| Redundancy | | 5.3 (5.1) |

**Table S2 Model refinement statistics**

| | | |
|---|---|---|
| Resolution range | | 92.80- 4.10 Å |
| Number of protein atoms | | 10296 |
| Number of solvent atoms | | None |
| Average B value | | 125.6 $Å^2$ |
| R-crystal (%) | | 17.9% |
| R-free | | 28.05% |
| Rms deviations | Bonds | 0.011 Å |
| | Angles | 1.721° |

Figure 1

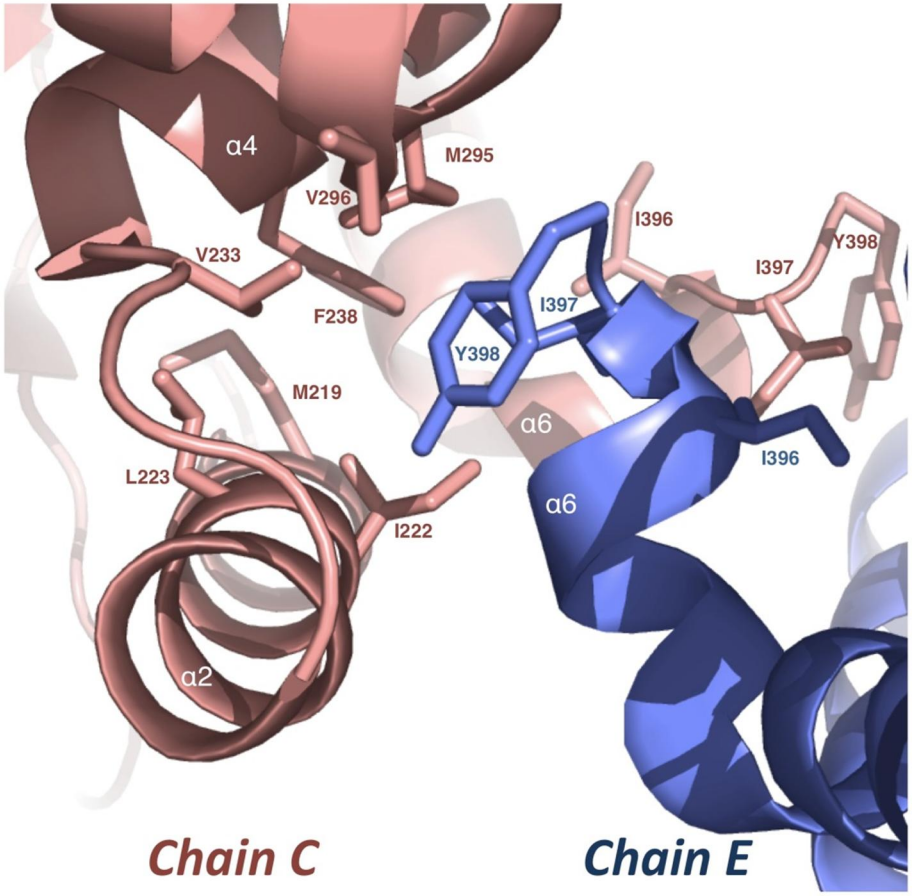

α4
M295
V296
V233
F238
M219
L223
I222
α2
I396
Y398
I397
I397
Y398
α6
I396
α6
Chain C
Chain E

Figure 2

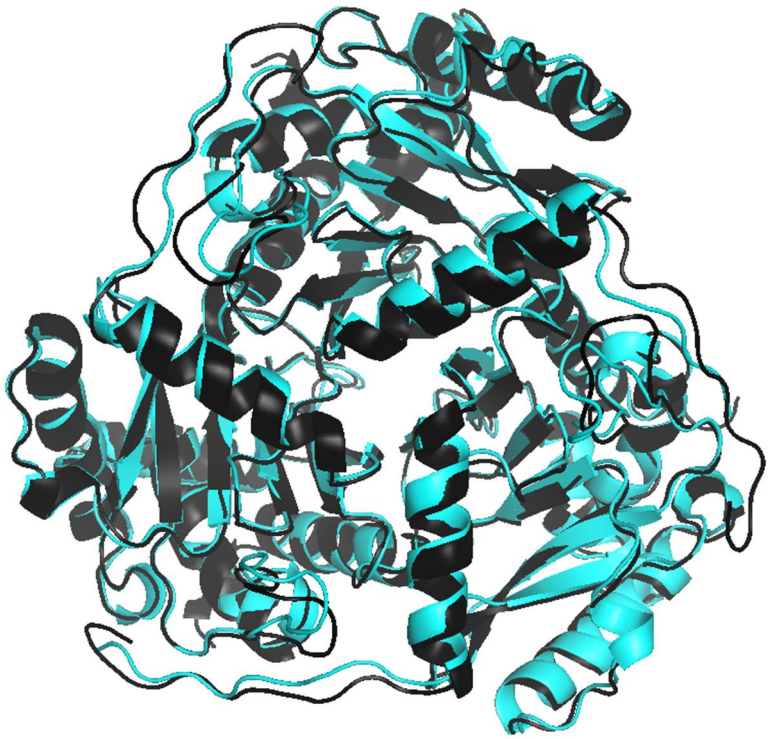


Figure 3

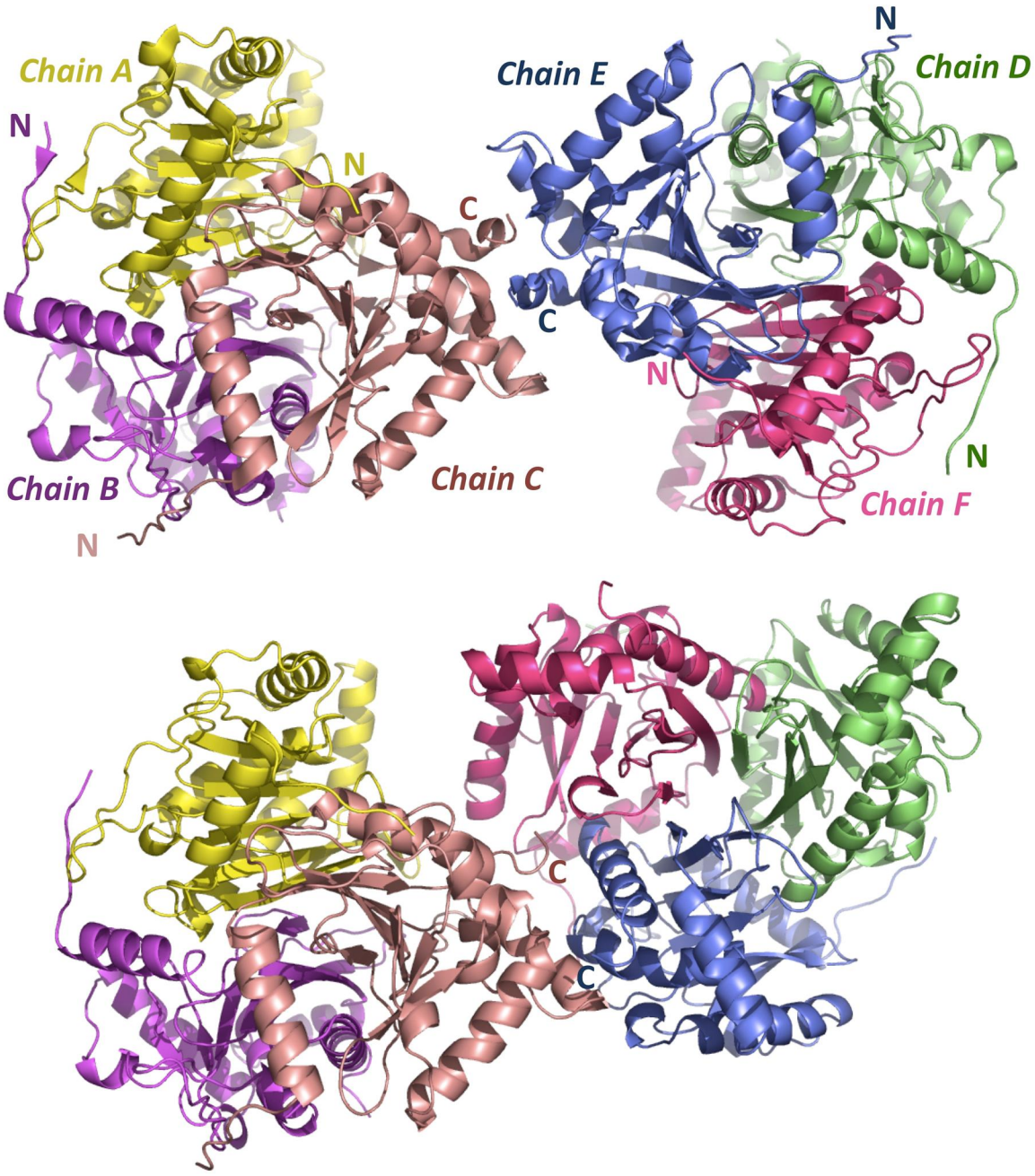


Figure 4a

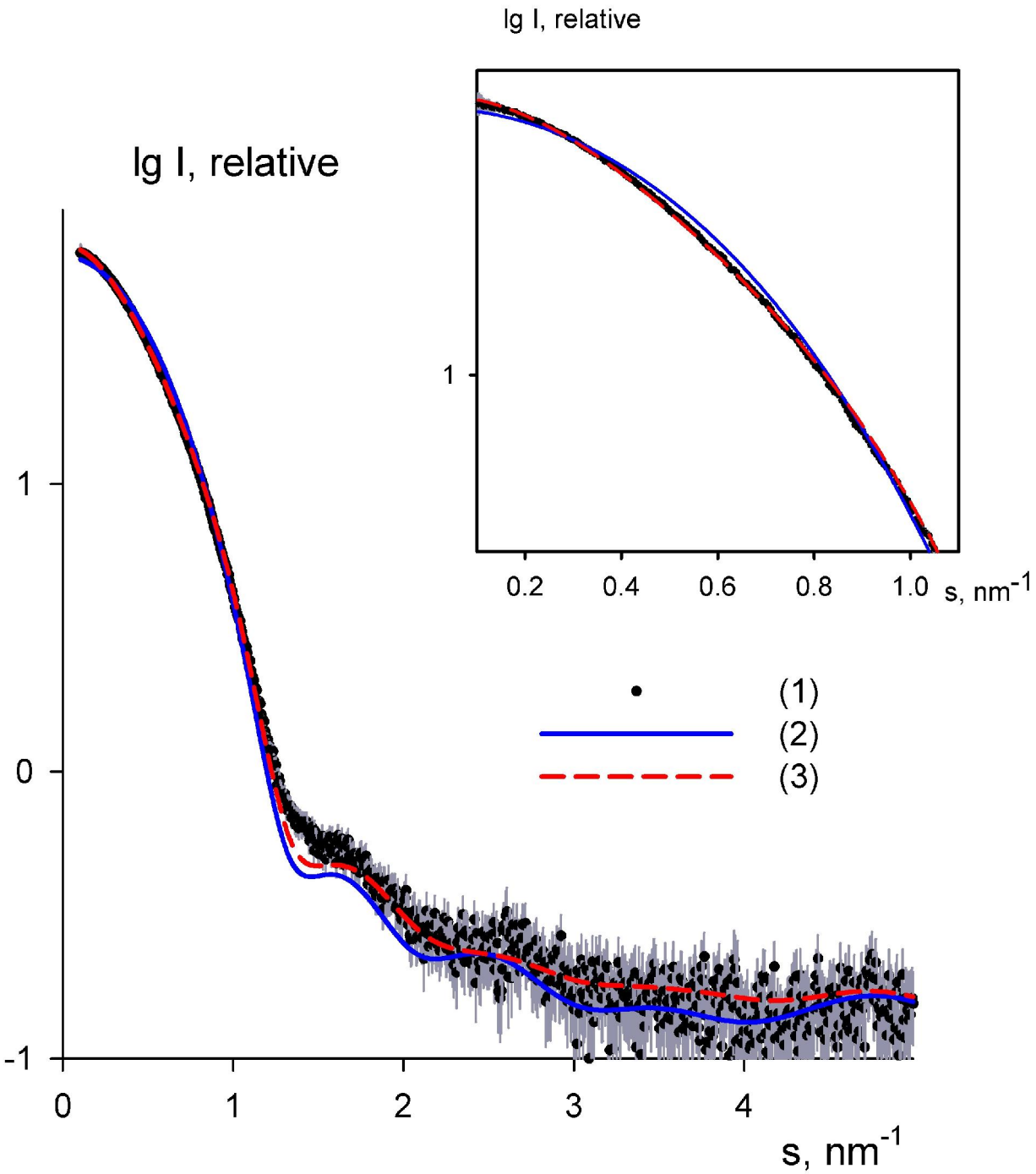

lg I, relative
lg I, relative
1
0.2
0.4
0.6
0.8
1.0
s, nm-1
1
0
-1
0
1
2
3
4
s, nm-1
(1)
(2)
(3)

Figure 4b

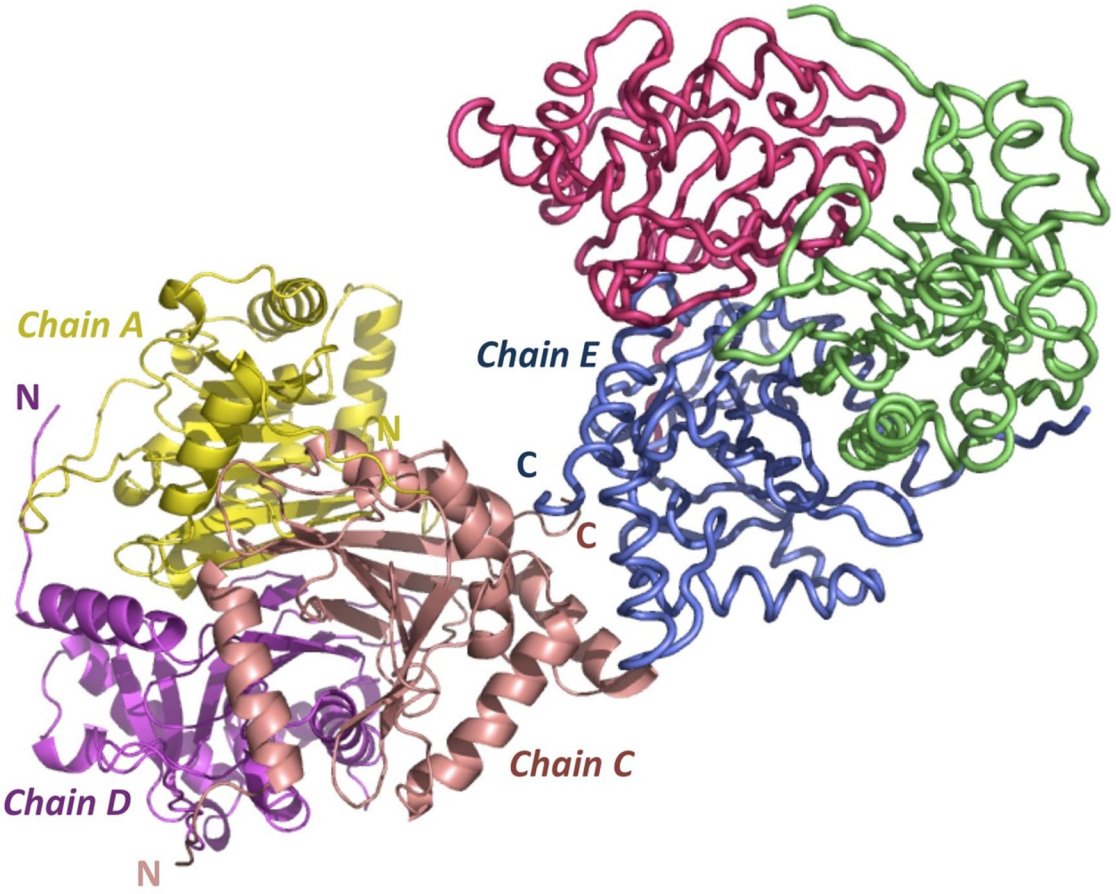


Figure 5

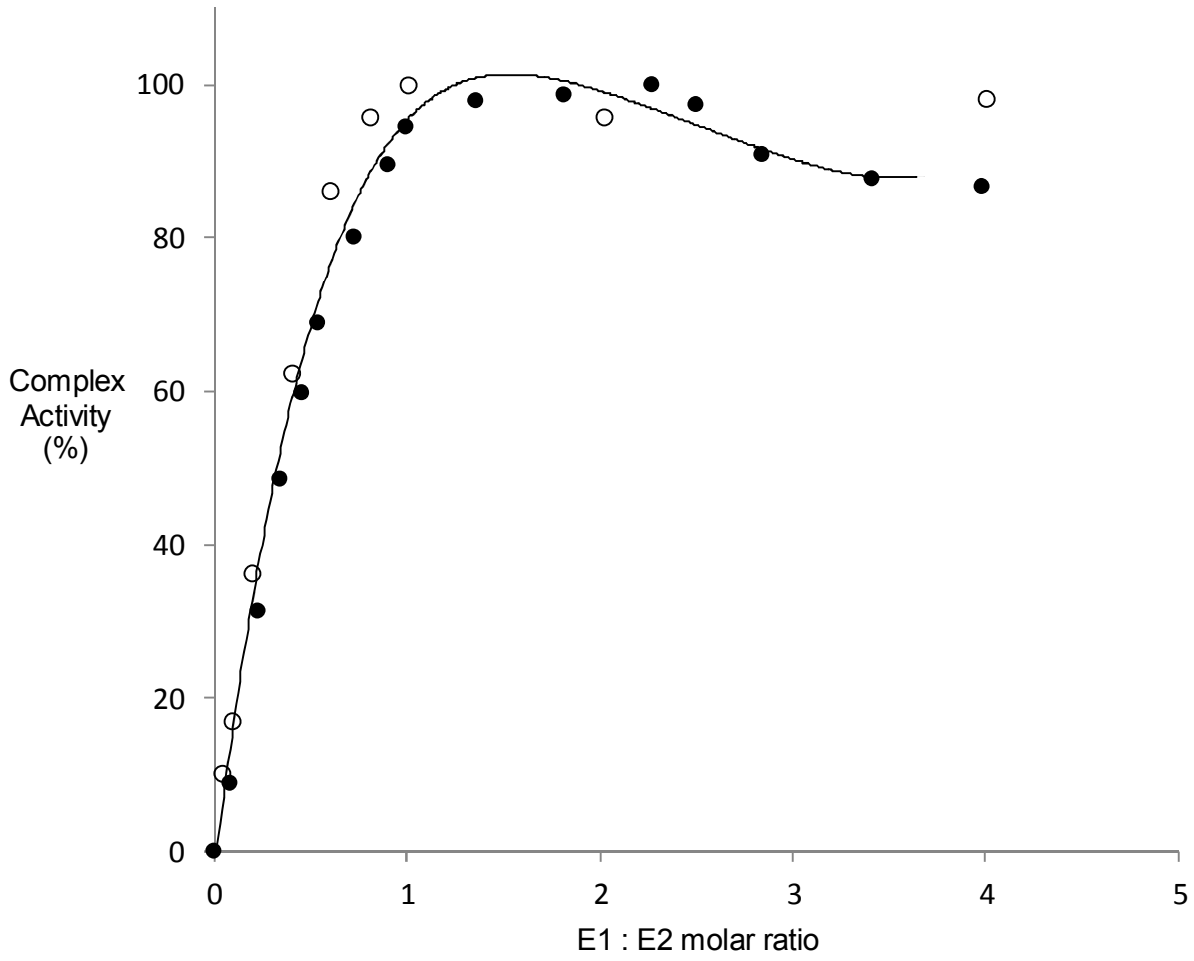
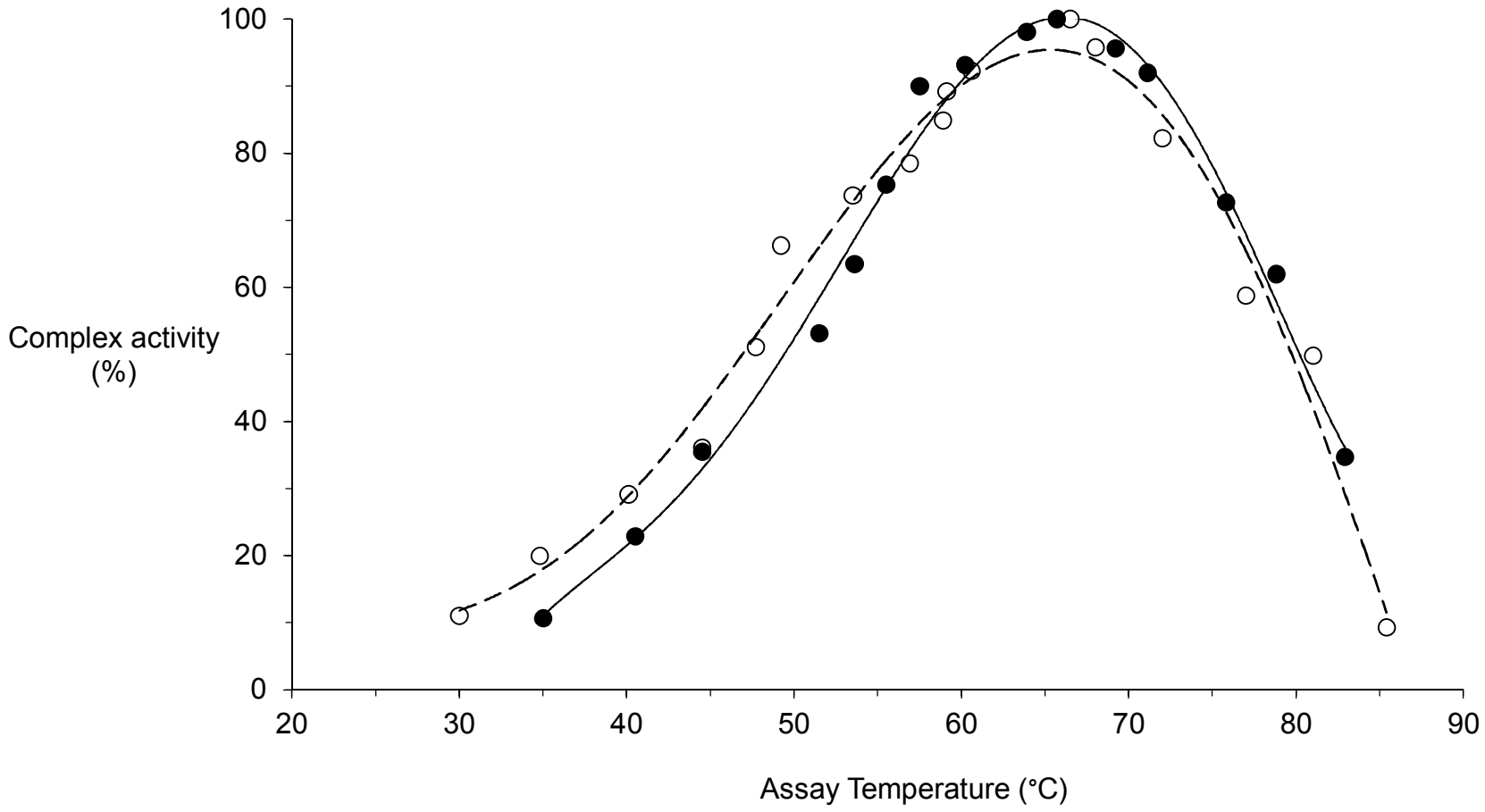


Figure 6


Complex activity
(%)
100
80
60
40
20
0
20
30
40
50
60
70
80
90
Assay Temperature (°C)