Physical Biology

OFFICIAL JOURNAL OF
Sociedad de Biofísicos Latino Americanos

ACCEPTED MANUSCRIPT

Salt bridge impact on global rigidity and thermostability in thermophilic citrate synthase

To cite this article before publication: Thomas James McManus *et al* 2019 *Phys. Biol.* in press https://doi.org/10.1088/1478-3975/ab2b5c

Salt bridge impact on global rigidity and thermostability in thermophilic citrate synthase

T J McManus[1], S A Wells[2], A B Walker[1]

[1] Department of Physics, University of Bath, Bath, BA2 7AY, UK
[2] Department of Chemical Engineering, University of Bath, Bath, BA2 7AY, UK

E-mail: T.J.McManus@bath.ac.uk

Abstract.

It has been suggested that structural rigidity is connected to thermostability, e.g. in enzymes from thermophilic microorganisms. We examine the importance of correctly handling salt bridges, and interactions which we term 'strong polars', when constructing the constraint network for global rigidity analysis in these systems. Through a comparison of rigidity in citrate synthases, we clarify the relationship between rigidity and thermostability. In particular, with our corrected handling of strong polar interactions, the difference in rigidity between mesophilic and thermophilic structures is detected more clearly than in previous studies. The increase in rigidity did not detract from the functional flexibility of the active site in all systems once their respective temperature range had been reached. We then examine the distribution of salt bridges in thermophiles that were previously unaccounted for in flexibility studies. We show that in hyperthermophiles these have stabilising roles in the active site; occuring in close proximity to key residues involved in catalysis and binding of the protein.

1. Introduction

Proteins are ubiquitous as structural and functional elements in living cells [1, 2, 3, 4]. A protein's function, for example as a catalyst (enzyme) or transporter (e.g. haemoglobin), is deeply tied to its structure. It has become evident that not only the static structure, but also the dynamics and flexibility of a protein are key to its function: many enzymes undergo substantial, reversible conformational changes in the course of their chemical operation. Since these dynamics are in turn a property of the protein structure and environment, there is substantial interest in simulation methods that can take as input a protein structure, that can be obtained through protein crystallography, and explore its motions on a wide variety of length and time scales. These range from the pico/nanosecond scale of small local rearrangements such as side chain reorientations, to the micro/millisecond scales of large domain motions such as hinge motions [5, 6].

The most physically detailed, and hence computationally expensive, simulation methods of classical molecular dynamics (MD) [7, 8], quantum-mechanical (QM) [9, 10], and hybrid QM/MM approaches, excel in detailed explorations of length and time scales of typically nm and ps respectively, but have difficulty reaching the micro/millisecond timescale without enormous expense in time and computational resources, substantial loss of detail through coarse-graining, or both. As a result, there is interest and value in alternative approaches for inferring useful information on rigidity and flexibility from structure input at lower computational expense [11, 12, 13]. For example, the nature of the lowest-frequency (slowest, and often largest) motions intrinsic to a protein structure can be inferred through elastic network modelling [14, 15].

Folding and unfolding of the protein, through which secondary/tertiary structures are formed from the primary structure (amino acid sequence), or broken up leading to a loss of functionality, is also of interest for a wide range of applications [16]. Of particular interest for the scope of this work, is the ability to simulate how the unfolding of an enzyme structure would take place, and how it is coupled with the stability of that protein. Previous work has suggested that structural rigidity is connected to thermostability, e.g. in enzymes from thermophilic microorganisms. We extend this investigation to examine how the connection between the two changes upon correctly classifying and handling salt bridges, and interactions which we term 'strong polars', when constructing the constraint network for global rigidity analysis effects. Through comparison with a well established method for exploring flexibility in protein structures 'FIRST' [10, 17, 18] we will demonstrate that these features are increasingly necessary to model in thermophilic species.

1.1. Extremozymes and Thermostability

Extremophiles are typically single-cell organisms that reside in extreme temperatures (below 25°C and above 50°C) and are commonly split into 4 categories. Psychrophiles (also known as cryophiles), mesophiles, thermophiles, and hyperthermophiles: with typical organism temperatures of 5-25°, 25-50°, 50-80°, and 80°+ respectively [19, 20]. The enzymes formed in these organisms tend to be stable, functional, and have their optimal activity in temperature ranges close to those of the organism, and so are named 'extremophilic enzymes' or 'extremozymes'. Many proteins exist as multiple homologous types in organisms across a wide range of temperatures. These types evolve for stability at their respective organism's temperature, so as to perform the same functional motion as one another - despite temperature differences of up to 100 degrees. One theory as to how this is achieved, is through an increased potency of electrostatic interactions contributing to rigidity at higher temperatures [21].

1.2. Calculating Rigidity

Rigidity and flexibility can be quantified in terms of degrees of freedom; independent parameters that define a mechanical system's configuration. A rigid body in three-dimensional space has 6 inherent degrees of freedom; 3 translational components defining position, and 3 rotational components defining orientation. A large molecular body, such as a protein, has many degrees of freedom; 6 of these describe position and orientation in space, while the remainder represent internal motions. Constraints within the structure will counterbalance the degrees of freedom. In the case of proteins these are distance constraints imposed by local geometry: covalent bonds, the hydrophobic effect, and electrostatic/ionic interactions such as hydrogen bonds and salt bridges. These distance constraints are expressed as edges within a network of atomic nodes, which subtract a degree of freedom from one of the nodes attached [22].

The pebble game algorithm - as implemented in the modelling software FIRST [23] - uses this principle, along with a method for rigidity percolation, to perform rigid cluster decomposition (RCD) of a protein expressed as a mechanical network in three dimensions as in Figure 1 (left). This procedure defines the flexibility of a protein by assessing the degrees of freedom of its constituent parts [23, 24, 25, 26, 27].

Figure 1 (right) displays a cluster decomposed structure overlapped with the locations of salt bridges in the enzyme. Whilst some of these salt bridges exist in rigid regions it is not as simple as one

salt bridge being enough to immediately rigidify its surrounding environment - hence the need for analysis using percolative theory.

We identify key transition points in the rigidity of thermophilic species in a manner similar to how hydrogen bond dilutions have been used to mimic thermal denaturation [20, 28, 29]. Our approach employs global rigidity measures as constraints are progressively removed in order of strength, from weakest to strongest, from a protein structure. We then compare the strengths of constraints that cause a large shift in rigidity in different structures. Relative measures, analagous to effective energies at which a structure will unfold or melt, are obtained across different species and used to discuss their relative thermodynamic stability.

1.3. Citrate Synthase

Citrate synthase, CS, is found in almost all living cells, catalyzing carbon entry in the citric acid cycle [30]. Its dimeric structure contains two subunits, each consisting of a large domain and a small domain. The two large domains come together to form the main bulk of the structure. The two hinges formed where the bulk meets the small domains mark the active sites. The sizes of the gaps created when the hinge is open, and the opening and closing motions of the hinges, vary between different CS structures. The functional hinge motion of the protein requires that the small domains not be mutually rigid with the central bulk of the enzyme. Due to being found in most organisms, CS can be found in environments ranging from as low as 0-10°C up to ∼100°C. As a result the more thermophilic structures have evolved a signature for higher rigidity, whilst maintaining the same functional motion. Understanding how this extremophilic stability is achieved without sacrificing function will lead to an increased understanding of protein stability as a whole, which will improve our ability to engineer proteins for application.

A previous study using FIRST [15] investigated the relationship between the need to stabilize the rigid structure of a protein and to maintain the necessary flexibility to perform its function, in proteins with large scale functional motions. By comparing thermal energy using the effective kcal/mol energy scale of FIRST, observations were made that in systems such as CS the 'folding core' of extremophilic proteins is more stable than those in mesophilic cases, with the rigidity of each protein corresponding to the temperature state of its organism. i.e. The rigidity of a thermophile at high temperature is approximately equal to that of a mesophile at room temperature. It is also expected that thermophiles and hyperthermophiles display a greatly increased rigidity over less thermophilic organisms when the analysis includes constraints whose strength is around the room temperature range of -0.5 to -1 kcal/mol; noting that at 300K the thermal energy k_BT is ∼0.6kcal/mol.

In this study we re-analyse a series of CS structures, making a correction to the handling of salt bridges and strong polar interactions, which were previously not properly accounted for [15]. This is essential given their importance to the rigidity of more thermophilic structures. The structures modelled are named using a prefix to demonstrate the organism in which they are found, and a numerical postfix for the state of their bound ligands. PigCS-0 for example, is a pig based structure with 0 ligands bound in the cleft of the active sites. Full details are given in Table 1.

2. Methods

We carry out rigidity analyses using the pebble game analysis implemented in the software FIRST [23]. We compare the rigidity analyses generated from FIRST's built-in constraint functions to those generated with our corrected handling of strong polar interaction energies. Our approach is implemented in a set of software tools which generate constraint files suitable for use as input to FIRST. These are available through the University of Bath research portal - SBFIRST dataset DOI: 10.15125/BATH-00566 [31]. All structures are taken from the PDB, hydrogenated with the web service Molprobity [32], and visualised with Pymol [33].

2.1. Hydrogen Bond and Salt Bridge Energies

The FIRST software makes use of a function to calculate an effective energy for polar interactions based on their geometry [34, 35], with terms based on the donor-hydrogen-acceptor (D-H-A) angle and on the donor-acceptor (D-A) distance (Figure 2). The latter takes the form of a Lennard-Jones (LJ) potential based on the D-A distance d, a distance correction term a, an equilibrium distance r_0, and a well depth D, as follows:

$$E_{LJ} = D\left(5\left(\frac{r_0}{d+a}\right)^{12} - 6\left(\frac{r_0}{d+a}\right)^{10}\right) \quad (1)$$

For hydrogen-bond interactions, values of $a = 0$, $r_0 = 2.8$, $D = -8$ kcal/mol are used, while for salt bridges (typically stronger, and assigned no dependence on the angular geometry) the values are $a = 0.375$, $r_0 = 3.2$, $D = -10$ kcal/mol. These expressions are based on the Dreiding potentials as used in the work of Mayo et al. [23, 34, 35, 36]. The distance correction term, a, gives a broader well for salt bridges, with a similar equilibrium separation to that of hydrogen bond.

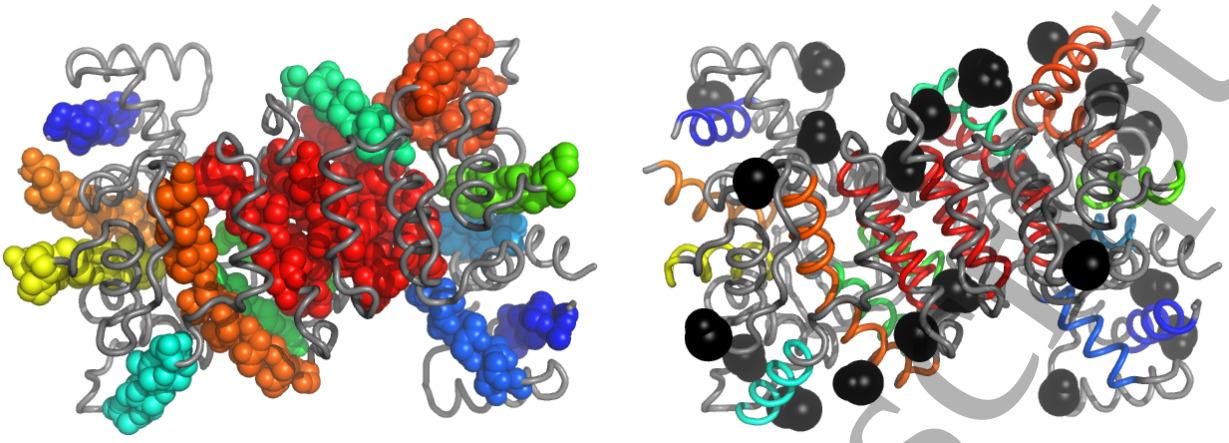

Figure 1: Left - A Rigid Cluster Decomposition (RCD) of *Thermus thermophilus* citrate synthase, including polar interactions stronger than (cutoff) kcal/mol. Each coloured sphere section represents a rigid cluster. A change in colour means that two adjacent sections are separate clusters to one another. Grey cartoon tube regions are completely flexible. Right - The same frame in ribbon structure with consistent colour and black spheres indicating the locations of salt bridges.

Prefix	Organism	Optimal (°C) Temperature	PDB (resolution Å) -0	-1	-2
AbCS	Arthrobacter DS2-3R	0-10	—	—	1a59 (2.09)
BsCS	Bacillus Subtilis	25	—	—	2c6x (3.40)‡
PigCS	Sus Scrofa	37	3enj (1.78)	—	2cts (2.00)
TaCS	Thermoplasma Acidophilum	59	—	2ifc (1.70)	2r9e (1.95)
TtCS	Thermus Thermophilus	70	1iom (1.50)	—	1ixe (2.30)
SsCS	Sulfolobus Solfataricus	80	1o7x (2.70)	—	—
PaCS	Pyrobaculum Aerophilum	100	—	2ibp (1.60)	—
PfCS	Pyrococcus Furiosus	100	—	—	1aj8 (1.90)

Table 1: The PDB codes and assigned prefix labels for the citrate synthase structures used in this study. ‡ - in the case of 2c6x the ligands are not bound to the functional state and the protein exists in its open form as an effective '-0' structure.

We have identified two difficulties with the implementation and interpretation of these functions in FIRST which may affect the accuracy of its analysis, especially for salt bridge interactions. Since such interactions are more common in thermophilic organisms as a stabilising mechanism, an accurate analysis is particularly relevant to the study of rigidity and thermostability. We discuss these difficulties and describe a corrected function, which we will term SBFIRST for the purposes of this paper. We provide software to read PDB files and produce hydrogen-bond energy lists, suitable for use as input to FIRST, using the SBFIRST function. This allows use of the FIRST rigidity analysis software with our corrected function.

The first difficulty is that, in the FIRST code base, the angular term includes an additional exponential pre-factor not found in the Dreiding forcefields and not discussed in any supporting literature. It appears that the effect of this term is generally negligible, typically only affecting bonds with near-90° D-H-A angles, which will in any case be weak. SBFIRST therefore omits this prefactor.

The more important difficulty concerns the interpretation of the energy function as used in FIRST, to calculate the energies of interactions observed in crystal structures, as opposed to its conventional interpretation in molecular simulations. For D-A distances less than the equilibrium value r_0, the LJ function is weakened and then becomes repulsive, with the bond energy increasing from the minimum (the well depth) and becoming less and less favourable, even becoming positive at sufficiently low D-A distances. In a molecular simulations context, this is a necessary and desirable feature. However, in the interpretation of protein crystal structure for rigidity analysis, we argue that this use of the LJ potential is inappropriate, on

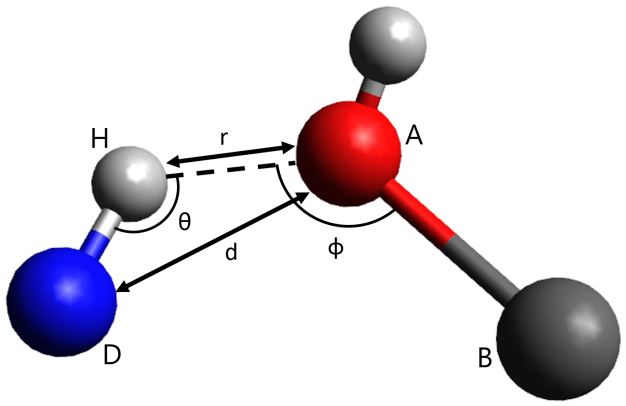

Figure 2: Definition of a hydrogen bond: Donor 'D' (blue), hydrogen 'H' (light grey), acceptor 'A' (red), and base 'B' (dark grey). H-A separation 'r', and D-A separation 'd' as used in Lennard-Jones potential term (equation 1). DHA angle 'θ ' and HAB angle 'ϕ ' from Dreiding equations and angular penalty terms.

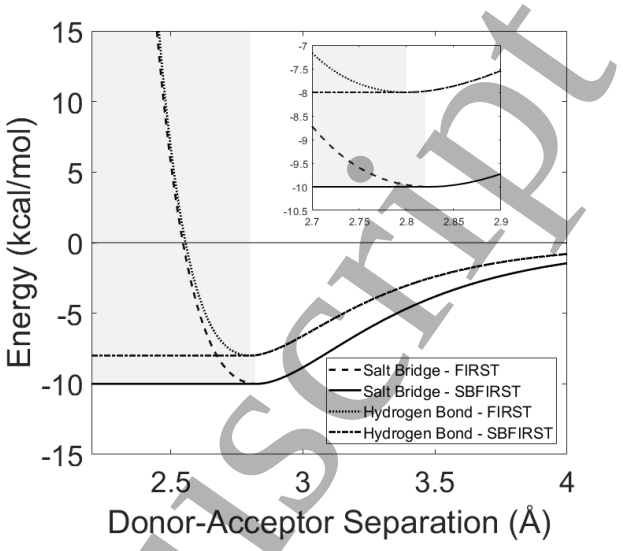

Figure 3: FIRST and SBFIRST Lennard-Jones potentials, showing corrected energy variation with Donor-Acceptor separation for interactions previously taking place in the weakened and repulsive (shaded) regimes at sufficiently low atomic separations.

both biological and physical grounds.

The biological argument is that, when a salt bridge with a short D-A distance (that is, a close approach of a positively and a negatively charged side group) is observed in a crystal structure, this is interpreted as a particularly strong and close interaction, particularly when studying thermophilic and hyperthermophilic systems [37, 38]. Assigning it a weak energy using the repulsive portion of the LJ potential does not match this biological interpretation.

The corresponding physical argument is that a protein in the cell is in a highly dissipative environment. Suppose that a salt bridge occurs in the structure with a short D-A distance, such that the interaction becomes weak or repulsive. As the structure relaxes, the D-A distance will move towards the equilibrium value, at which the energy of the interaction is at a minimum (the well depth). Energy released in the process will be dissipated through the many vibrational modes of the protein and ultimately into the thermal bath of the solvent/cellular environment surrounding it.

On both of these grounds, therefore, the SBFIRST function does not make use of the negative gradient portion of the LJ potential. Rather, all interactions with D-A distances less than r_0 are assigned the energy corresponding to the minimum of the potential as in Figure 3. This ensures that close salt bridge interactions are always interpreted as favourable.

To detect hydrogen bonds and salt bridges throughout a protein SBFIRST makes use of an in house structural parser, which populates an array of all polar atoms based on the local chemistry of each site. In doing so searches are also conducted for guanidino groups, ionized carboxyl groups, and imidazolate groups; as commonly found in arginine, aspartamine and glutamic acid, and histidine residues respectively. The (de)protonation state of these groups determines their affinity for polar and ionic constraint formation.

2.2. Rigidity Fraction

The rigidity fraction [39], RF, uses the pebble game algorithm (section 1.2) to quantitatively calculate the rigidity of a protein as a function of an effective energy scale. The hydrophobic effect is always included as constraints in the protein's network, although the strength of these constraints is reduced in the case of psychrophilic enzymes to account for the reduced potency of entropic effects in these systems. Electrostatic effects are assigned an effective energy in the range of -10 to 0 kcal/mol. Each constraint is then only included in the RCD if its energy lies below a given cut-off value. After RCD the ith RF, RF_i, measures the proportion of α-carbon atoms found in the first i largest rigid clusters. By progressively lowering the energy cut-off parameter across the energy range 0.0 to -4.0 kcal/mol, relative rigidities are obtained for the system at different effective energies. The retention of rigidity as the cut-off is lowered is a signal of greater structural rigidity, since the structure would be more resistant to the bond breaking effects of a hotter thermal bath than a structure which loses rigidity under the same conditions.

3. Results and Discussion

3.1. Comparative Rigidity

Figure 4 shows an RF_5 analysis of CS from 6 different species; the mesophilic reference organism, PigCS, and 5 species of varying extremophilic nature using files of higher resolution ($\sim$ 2.00 Å) when possible. This is done using FIRST's built-in energy functional (top), our corrected functional SBFIRST (middle), and the change in rigidity, ΔRF, between the two (bottom). The inclusion of our correction for the salt bridge energies displaces the rigidity fraction to a higher value in all cases - as would be expected with the introduction of additional constraints into the system. Relative rigidity of the different species remains unchanged; thermophiles still retain rigidity in the bulk of the protein for longer than other organisms, and mesophiles for longer than psychrophiles. Both mesophiles and thermophiles exhibit a greatly increased rigidity over psychrophiles in the -0.5 to -1 kcal/mol range, corresponding to room temperature ($\sim$0.6 kcal/mol), which is known to be a key aspect of the relationship between rigidity and thermostability.

The psycrophilic ABCS-2 enzyme and mesophilic PigCS-0 enzyme both show a sharp increase in ΔRF at the energetically low -0.5 kcal/mol cut-off; 0.27 and 0.19 respectively. This sharp increase however drops off immediately and has no impact on rigidity fraction in the more negative energy ranges. PfCS-2, PaCS-1, and TaCS-0 already exhibit the highest rigidity without the correction for salt bridge interactions, and are three of the most thermophilic examples studied - with organism temperatures of 100°C, 100°C, and 59°C respectively. The increased rigidity in these thermophilic species is observed at more negative energy cut-offs, corresponding to higher effective energies.

The details of the change in rigidity vary in their exact nature for each thermophilic species. PaCS-1 (one of the two 100°C hyperthermophiles in this study) exhibits a ΔRF of on average 0.17 across a broad range of cut-off values from -1.0 to -3.0 kcal/mol. TaCS-1 has only one point of high ΔRF at the -2 kcal/mol point in the effective energy scale: an increase in rigidity fraction of 0.49. PfCS-2 displays its increased rigidity at the -1.0 and -1.5 kcal/mol cut-offs..

Figure 5 shows RCD images of CS from 4 different organisms across the temperature spectrum - BsCs-2 (25°C), PigCS-2 (37°C), TtCS-2 (70°C), PaCS-1 (100°C) - at 4 cut off values in the effective energy scale. The constraint analysis leading to these images was all performed with the SBFIRST method, as to correctly account for the salt bridges previously discussed. For any one cut off value it can be observed that the higher the temperature of origin of the enzyme, the more rigid it appears when compared to enzymes of a lower temperature. Functional flexibility is attained at the point when the small domains are not mutually rigid with the bulk. This behaviour is highlighted where each diagonal term is surrounded by a box; each exhibiting a similar level of functional-rigidity, and confirming that a thermophilic folding core will retain rigidity corresponding to its organism's temperature. Despite this increased stability, once the functional-rigidity is obtained the active site remains equaly flexible across all different levels of thermophilia.

Correct interpretation of the prevalent salt bridges in thermophilic species does not therefore affect the retained functional flexibility of the protein, only the effective energy at which the bulk of the protein gains flexibility. The relative rigidity of each species also remain intact, and it can be observed that salt bridges have a higher impact on the rigidity of proteins, the more thermophilic the environment of their organism.

3.2. Newly Detected Salt Bridges

Figure 6 shows the location of the salt bridge detected by SBFIRST but omitted by FIRST in PfCS-2, present symmetrically in each monomer. In close proximity to two citrate binding residues, arginine residue 356 and glutamic acid residue 189, this stabilising interaction is immersed in the active site. The omission of this interaction would clearly lead to inaccurate handling of the geometry of active site residues involved in the proteins function.

Figure 7 (left) displays a number of salt bridges previously unaccounted for in the SsCS-0 structure. One of particular interest (Figure 7 right), exists between two α-helices in the active site (helices 15 and 16 in the pdb file). Relative chain alignment of neighbouring helices is known to be governed by inter-helical salt bridges in α-helix domains. Multiple catalysis, citrate binding, and phosphate binding sites exist either as part of the two bound helices or in their neighbouring loops. Their orientation and alignment in a configurational study would therefore be unreliable in the absence of the domain defining salt bridges now found to be present in the flexibility modelling suite SBFIRST.

4. Conclusion

In order to examine the effect of correctly handling salt bridges when examining rigidity and thermostability, a new energy functional termed 'SBFIRST' was discussed using the existing method implemented in 'FIRST' as a control. This method altered the potential used when assessing polar interactions within a molecular structure, to strengthen and in

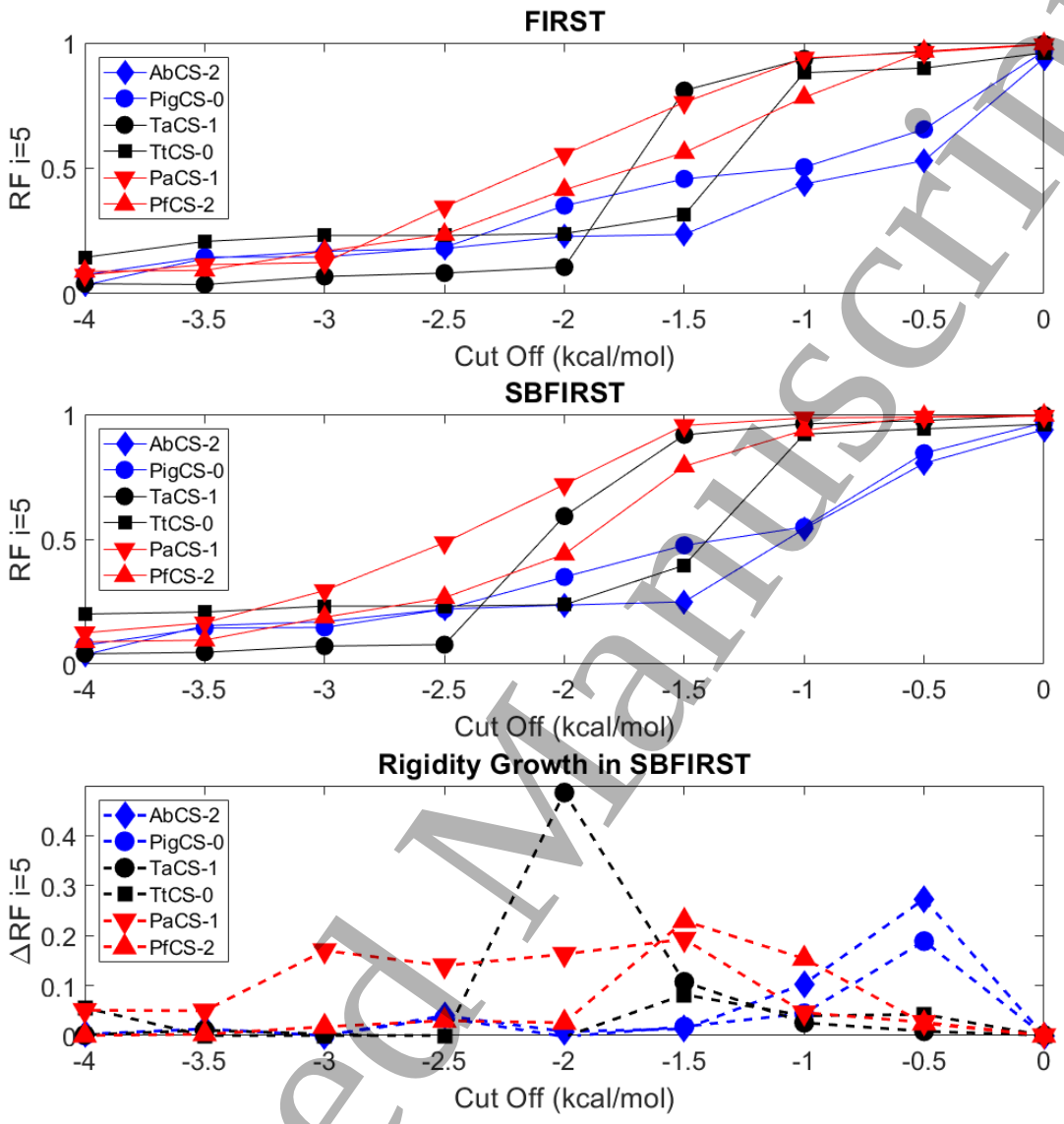

Figure 4: FIRST (top) and SBFIRST (middle) rigidity fraction, $i = 5$, analysis of a range of thermophilic citrate synthase structures, and the growth of rigidity fraction in SBFIRST compared to FIRST (ΔRF) (bottom). Psychro/mesophiles represented in blue, thermophiles in black, hyperthermophiles in red.

some instances include previously unnacounted-for salt bridges due to their importance in stabilising thermophilic organisms. This potential was used as an effective energy scale through FIRST in a percolative rigidity analysis, in order to observe impact on the simulated denaturation of the folding core of (hyper)thermophilic, mesophilic, and psychrophilic CS.

A comparison of rigidity in thermostable CS both with and without correct salt bridge assessment in the structure showed, whilst clarifying the relationships between species of varying thermophilic natures, that rigidity was increased by an overall factor. This factor increased in organisms of a higher thermophilic nature, due to the abundance of strong salt bridge interactions within their molecular structure. Psychrophilic and mesophilic enzymes exhibited increased rigidity at the smallest cut-off values of 0 to -1 kcal/mol and not beyond. The exact pattern of thermophilic rigidification varied in detail between structures, with

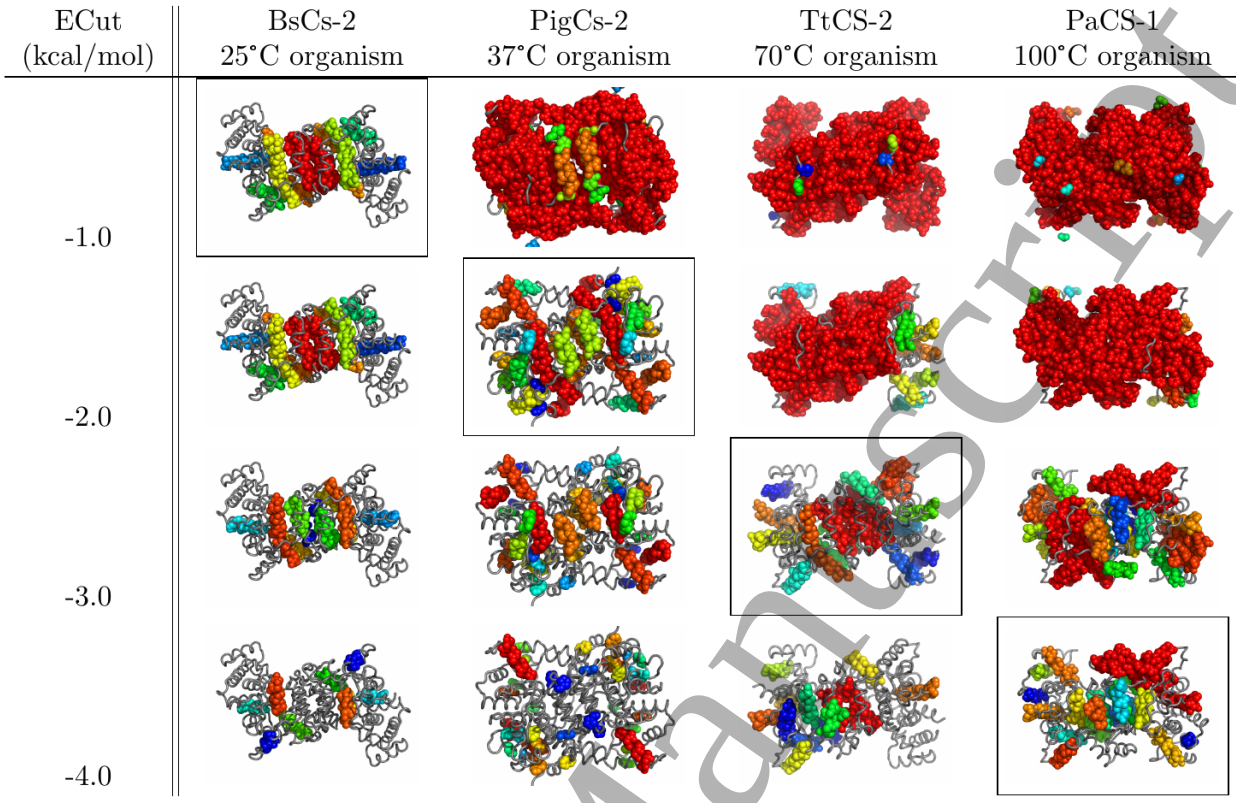

Figure 5: Rigid cluster decomposition of Citrate Synthase species across the thermophilic spectrum at constraint inclusion cut-offs from -1 to -4 kcal/mol.

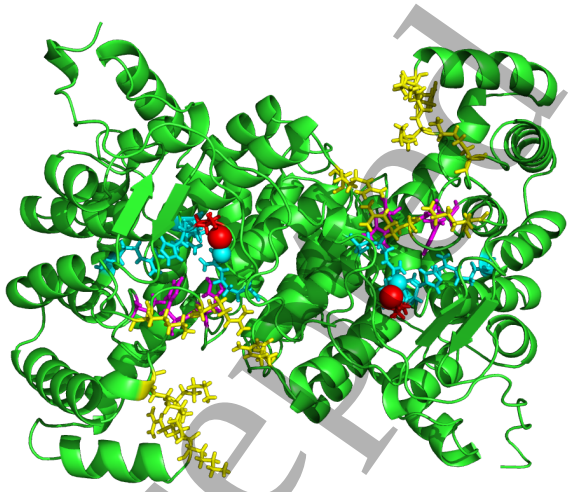

Figure 6: Newly detected salt bridges in PfCS-2 by SBFIRST. Salt bridges (red sticks with spheres), catalytic residues (magenta sticks), citrate binding residues (cyan sticks) and phosphate binding residues (yellow sticks).

for example TaCS displaying dramatic rigidification at the -2.5 cal/mol cutoff specifically, whereas PfCS was rigidified at cutoffs of -1 to -1.5 kcal/mol, and PaCS shows greater rigidity across a wide range of cutoffs from -1.5 to -3 kcal/mol. The increase in rigidity did not affect the functional flexibility of the active site in all systems once their respective temperature range had been reached.

SBFIRST is found in thermophiles and hyperthermophiles to detect salt bridges that were either missing entirely, or assigned a positive repulsive energy, in the initial FIRST method. In thermophiles such as TaCS-2 these exist typically as interchain bridges present in the central bulk of the dimer. In hyperthermophiles SsCS-0 and PfCS-2 these are observed in the active site stabilising the positions of key binding and catalysis sites.

We anticipate that the greater clarity offered by this handling of salt-bridge and strong polar interactions will be generally useful for protein rigidity analysis.

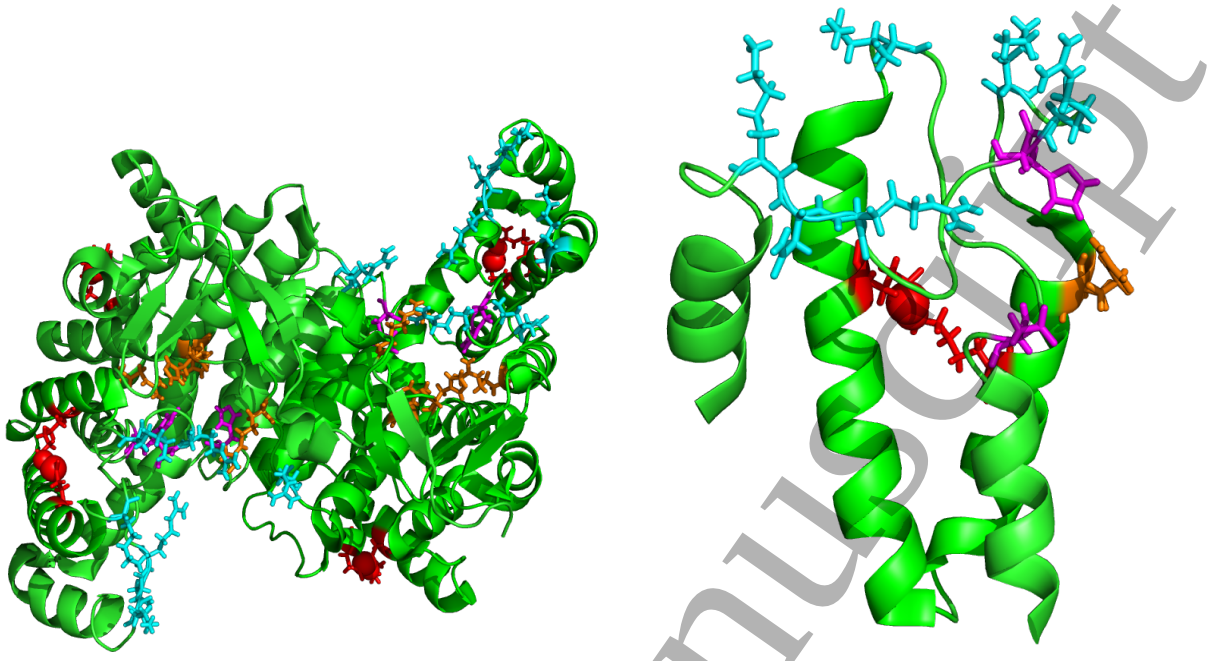

Figure 7: Newly detected salt bridges in SsCS-0 by SBFIRST. Salt bridges (red sticks with spheres), catalytic residues (magenta sticks), citrate binding residues (orange sticks) and phosphate binding residues (cyan sticks). Whole structure (left) and close up of helices 15 and 16 in the active site (right).

Acknowledgements

TJM was funded jointly by a studentship from the University of Bath, and ClusterVision's HPC Scholarship Programme. SAW has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648283 "GROWMOF").

[1] Murphy I 1995 *Theory and Practice* **3** 109–115
[2] Berthoumieu O, Patil A V, Xi W, Aslimovska L, Davis J J and Watts A
[3] Hampp N 2000 *Chemical reviews* **100** 1755–1776 ISSN 1520-6890
[4] Kato H E, Kamiya M, Sugo S, Ito J, Taniguchi R, Orito A, Hirata K, Inutsuka A, Yamanaka A, Maturana A D, Ishitani R, Sudo Y, Hayashi S and Nureki O 2015 *Nature communications* **6** 7177 ISSN 2041-1723
[5] Oleinikovas V, Saladino G, Cossins B P and Gervasio F L 2016 *Journal of the American Chemical Society* **138** 14257–14263 ISSN 15205126
[6] Touw W G, Joosten R P and Vriend G 2016 *Journal of Molecular Biology* **428** 1375–1393 ISSN 10898638
[7] Giorgetti A and Carloni P 2014 Protein Modelling pp 165–174 ISBN 978-3-319-09975-0
[8] Russo J, Tavares J M, Teixeira P I C, Telo Da Gama M M and Sciortino F 2011 *Physical Review Letters* **106** 1–4 ISSN 00319007
[9] Náray-szabó G 2014 Protein Modelling pp 1–4 ISBN 978-3-319-09975-0
[10] Römer R A, Wells S A, Emilio Jimenez-Roldan J, Bhattacharyya M, Vishweshwara S and Freedman R B 2016 *Proteins: Structure, Function and Bioinformatics* **84** 1776–1785 ISSN 10970134
[11] Fuentes G, Dastidar S G, Madhumalar A and Verma C S 2011 *Drug Development Research* **72** 26–35 ISSN 02724391
[12] Hernández G, Anderson J S and Lemaster D M 2012 *Protein NMR Techniques* vol 831 ISBN 978-1-61779-479-7
[13] Nussinov R 2001 *IBM Journal of Research and Development* **45** 499–512
[14] Suhre K and Sanejouand Y H 2004 *Nucleic Acids Research* **32** 610–614 ISSN 03051048
[15] Wells S A, Crennell S J and Danson M J 2014 *Proteins: Structure, Function and Bioinformatics* **82** 2657–2670 ISSN 10970134
[16] Zheng W M 2014 *Chinese Physics B* **23** ISSN 16741056 (*Preprint* `arXiv:1311.6762v1`)
[17] Wells S A, Van Der Kamp M W, McGeagh J D and Mulholland A J 2015 *PLoS ONE* **10** ISSN 19326203
[18] Gohlke H and Thorpe M F 2006 *Biophysical Journal* **91** 2115–2120 ISSN 00063495
[19] Vieille C and Zeikus G J 2001 *Microbiology and Molecular Biology Reviews* **65** 1–43 ISSN 1092-2172
[20] Rader A J 2009 *Physical Biology* **7** 016002 ISSN 1478-3975
[21] Karshikoff A, Nilsson L and Ladenstein R 2015 Rigidity versus flexibility: The dilemma of understanding protein thermal stability
[22] Hale L C 1999 *Dissertation*
[23] Jacobs D J, Rader A J, Kuhn L A and Thorpe M F 2001 *Proteins* **44** 150–65 ISSN 0887-3585
[24] Jacobs D J and Thorpe M F 1995 *Physical Review Letters* **75** 4051–4054 ISSN 00319007
[25] Thorpe M, Jacobs D and Djordjevic B 1996 Generic Rigidity Percolation *Condensed Matter Theories* vol 11 pp 407–424
[26] Jacobs D J and Hendrickson B 1997 *Journal of Computational Physics* **137** 346–365
[27] Jacobs D J 1998 *Journal of Physics A: Mathematical and General* **31** 6653–6668

[28] Hespenheide B M, Rader A J, Thorpe M F and Kuhn L A 2002 *Journal of Molecular Graphics and Modelling* **21** 195–207 ISSN 10933263

[29] Chubynsky M, Hespenheide B M, Jacobs D J, Kuhn L a, Lei M, Menor S, Rader A, Thorpe M F, Whiteley W and Zavodszky M I 2008 *Nanotechnology Research Journal* **2** 61–72

[30] Wiegand G and Remington S J 1986 *Annual Review of Biophysics and Biophysical Chemistry* 97–117 ISSN 0883-9182 (*Preprint* `arXiv:1505.02419v2`)

[31] McManus T J and Wells S A 2018 Dataset for "Salt bridge impact on global rigidity and thermostability in thermophilic citrate synthase"

[32] Davis I W, Leaver-Fay A, Chen V B, Block J N, Kapral G J, Wang X, Murray L W, Arendall III W B, Snoeyink J, Richardson J S and Richardson D C 2007 *Nucleic Acids Research* **35**

[33] Schrödinger, LLC 2015 The PyMOL molecular graphics system, version 1.8

[34] Mayo S L, Olafson B D and Iii W a G 1990 *Journal of Physical chemistry* **101** 8897–8909 ISSN 0022-3654

[35] Dahiyat B I, Gordon D B and Mayo S L 1997 *Protein science : a publication of the Protein Society* **6** 1333–7 ISSN 0961-8368

[36] Lim T C 2004 *Journal of Mathematical Chemistry* **36** 261–269 ISSN 02599791

[37] Gong H and Freed K F 2010 *Biophysical Journal* **98** 470–477 ISSN 15420086

[38] Elcock A H 1998 *Journal of Molecular Biology* **284** 489–502 ISSN 00222836

[39] Wells S, Jimenez-Roldan J and Römer R 2009 **046005** ISSN 1478-3975 (*Preprint* `0905.4232`)