CHAPTER 6.A

The active site in a single-chain enzyme

Mateusz Banach[1], **Leszek Konieczny**[2], **Irena Roterman**[1]
[1]Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Krakow, Poland
[2]Chair of Medical Biochemistry, Jagiellonian University – Medical College, Krakow, Poland

Contents

The image visualizes a single-chain enzyme which, while largely accordant with the 3D Gaussian distribution, also contains a binding cavity. In terms of the fuzzy oil drop model the cavity manifests itself as a localized hydrophobicity deficiency. It encodes information which determines the function of the protein.

From Globular Proteins to Amyloids
ISBN: 978-0-08-102981-7
https://doi.org/10.1016/B978-0-08-102981-7.00007-5

A single-chain enzyme should—in accordance with information theory—encode information enabling it to recognize the intended substrate and perform catalysis. Its structure should be shaped by the aqueous environment (i.e. it should resemble a spherical micelle) and feature a local deficit of hydrophobicity in the area of its binding cavity. The deficit expresses information which determines the enzyme's biological role. We will investigate this hypothesis by considering two distinct single-chain enzymes: a lysozyme and a ribonuclease.

In this work the hydrolase family is represented by human lysozyme (EC 3.2.1.17, PDB ID: 1:Z1) [1] and *Bos taurus* pancreatic ribonuclease (EC.3.1.27.5, PDB ID: 5RSA) [2]. Both have a diverse secondary structure with multiple disulphide bonds and a clearly distinguished catalytic active site.

Lysozyme

Table 6.A.1 lists the structural properties of the lysozyme (PDB ID: 1LZ1). In light of the presented results, the lysozyme can be said to deviate from the theoretical distribution of hydrophobicity. Its RD (T-O-R) is greater than 0.5, with balanced values of all correlation coefficients (although HvO dominates, high TvO indicates strong involvement of hydrophobic forces in shaping the protein's structure).

Regarding catalytic residues and their immediate neighborhood, these fragments are also identified as discordant. Eliminating these residues lowers the RD(T-O-R) value for the remainder of the protein.

Fig. 6.A.1. Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for 1LZ1. Magenta background denotes helical fragment while yellow background denotes β-strands. Orange stars mark the catalytic residues. Orange lines correspond to SS-bonds formed by Cys residues.

The structure of the lysozyme is further stabilized by four disulfide bonds. Fragments bracketed by these bonds vary in terms of their FOD status: 6—128 and 30—116 are both identified as discordant while the other two fragments remain consistent with the model. Notably, the discordant fragments (6—128 and 30—116) contain both of the enzyme's catalytic residues. It is therefore possible that their discordance is not directly caused by the presence of disulfides.

When analyzing individual secondary folds, we note that the β-sheet as a whole, as well as all its component β-strands, significantly diverge from the

Table 6.A.1 Values of fuzzy oil drop parameters calculated for the structure of 1LZ1 and its selected fragments. Values printed in boldface reflect discordance. The leftmost column lists secondary folds and (where applicable) the location of catalytic residues.

Lysozyme (1LZ1)		RD		Correlation coefficient		
Fragment	**AA**	**T-O-R**	**T-O-H**	**HvT**	**TvO**	**HvO**
Complete mol.	1–130	**0.529**	0.477	0.387	0.515	0.781
Catalytic active site						
Cat. Res. absent	35, 53	**0.515**	0.465	0.432	0.547	0.779
35E	30–41	**0.504**	0.170	0.088	0.393	0.701
Close neighborhood ±5 residues						
53D	48–59	**0.545**	**0.649**	0.541	0.687	0.791
Close neighborhood ±5 residues						
SS-bonds						
SS-bonds	6–128	**0.531**	**0.504**	0.400	0.507	0.789
35E, 53D	30–116	**0.543**	0.462	0.387	0.499	0.762
	65–81	0.279	0.279	0.608	0.836	0.822
	77–95	0.488	**0.734**	0.365	0.493	0.832
Secondary structure						
β-strands	42–46	**0.620**	**0.750**	-0.194	0.345	0.525
53D	51–55	**0.603**	**0.626**	0.088	0.421	0.854
	59–61	**0.858**	0.428	0.802	-0.416	0.207
β-sheet		**0.617**	**0.514**	0.292	0.442	0.462
Helices	4–15	0.376	0.398	0.573	0.661	0.912
35E	24–37	**0.565**	0.213	0.190	0.237	0.642
	81–86	0.443	**0.779**	0.533	0.458	0.881
	89–101	0.263	0.118	0.562	0.778	0.813
	104–109	**0.502**	**0.612**	0.896	0.440	0.767
	109–116	0.317	**0.626**	0.637	0.696	0.899
	121–125	0.129	**0.622**	0.920	0.966	0.961

theoretical distribution of hydrophobicity. This is an important observation given the presence of the catalytic residue at 53D. Likewise, the helical fold which contains the enzyme's other catalytic residue (35E) is also discordant.

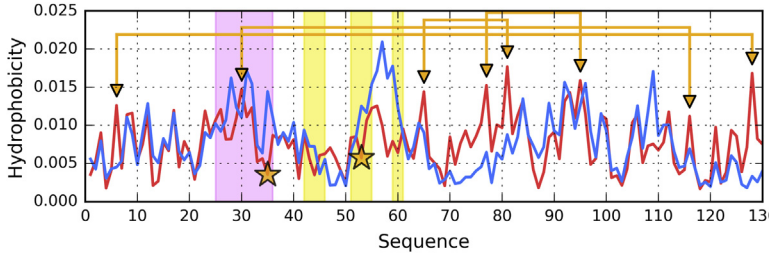

Fig. 6.A.1 Illustrates the status of the lysozyme, highlighting fragments which are identified as discordant in Table 6.A.1. As shown, all these fragments are close to the active site.

By invoking our formal hypothesis, which conflates information content with departures from the deterministic distribution of hydrophobicity, we may claim that the β-sheet and the neighboring helix carry information required by the enzyme to fulfill its role.

Fig. 6.A.2 clearly shows an outer shell which surrounds the active site. This shell exhibits micellar properties and is generated via interaction with the aqueous solvent. As already noted, information content determines activity—this includes the capability to recognize the substrate, as well as a blueprint for conformational changes required in the process of catalysis (which is inherently dynamic). It seems that the latter property may be linked to the presence of a centrally placed helix which plays an important part in stabilizing the molecule as a whole.

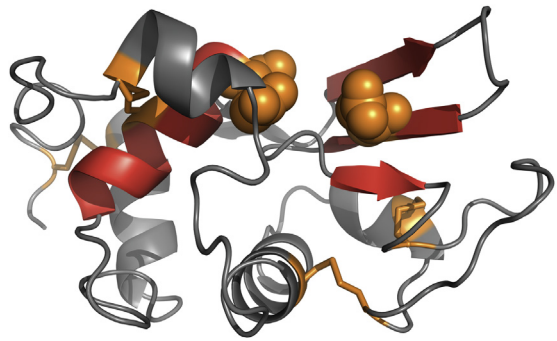

Fig. 6.A.2 3D presentation of 1LZ1. Fragments highlighted in red are identified by fuzzy oil drop model as discordant (helix: 24—37, sheet: 42—46, 51—55, 59—61). Orange spheres mark the catalytic residues (53D and 35E). Orange sticks correspond to disulfides.

Referring again to the status of the β-sheet, the status of its individual folds varies. The fragment at 51—55 is significantly affected by intrinsic hydrophobicity and strongly discordant—as evidenced by its high value of RD. The fragment at 42—46 is characterized by negative correlation coefficients coupled with high values of RD—we may describe it as being "in active opposition" to the theoretical distribution. The fragment therefore contributes a large quantity of information to the β-sheet and the entire neighborhood of the active site.

Ribonuclease

Ribonuclease (PDB ID: 5RSA) provides another example of how information can be encoded in a single-chain enzyme. Similarly to the lysozyme, we expect the protein to include a region characterized by strong discordance from the theoretical distribution, reflecting its high information content (Fig. 6.A.3 and Fig. 6.A.4).

It should be noted that fragments which comprise and surround the catalytic active site are non-micellar in character. These fragments are believed to encode information which determines the enzyme's specific activity profile.

As can be seen in Table 6.A.2, the discordance of the additional white fragment (87—90) identified in the presented hydrophobicity profiles (Fig. 6.A.4) may be related to the flexibility of the outlying loop which provides a way for the substrate to migrate to its required location.

The structure of ribonuclease exhibits major deviations from the theoretical hydrophobicity profile in areas which comprise its active site. Much like

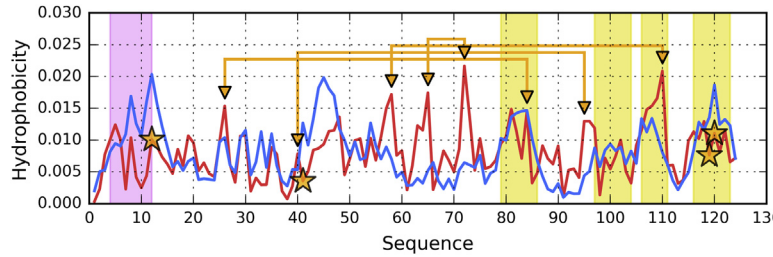

Fig. 6.A.3 Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for 5RSA. Magenta background denotes helical fragment while yellow background denotes β-sheets. Orange stars mark the catalytic residues. Orange lines correspond to SS bonds formed by Cys residues.

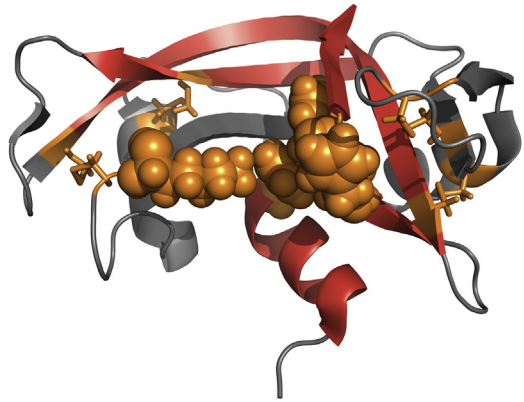

Fig. 6.A.4 3D presentation of ribonuclease (5RSA). Fragments highlighted in red are identified by fuzzy oil drop model as discordant (helix: 24—37, sheet: 42—46, 51—55, 59—61). Orange spheres mark the catalytic residues (53D and 35E). Orange sticks correspond to disulfides.

Table 6.A.2 Values of fuzzy oil drop parameters calculated for the structure of ribonuclease (5RSA) and its selected fragments. Values printed in boldface reflect discordance expressed either by high RD values or biased domination of HvO correlation coefficient. The leftmost column lists secondary folds and (where applicable) the location of catalytic residues. Underscored values are regarded as particularly important for determination of conformational characteristics.

		RD		Correlation coefficient		
Ribonuclease (5RSA) Fragment	AA	T-O-R	T-O-H	HvT	TvO	HvO
Complete mol.	1—124	**0.550**	**0.563**	0.300	0.378	0.815
Catalytic center						
Close neighborhood ±5 residues						
12H	7—17	0.469	0.315	0.385	0.548	0.837
41K	36—46	0.407	0.281	0.278	0.647	0.820
119H, 120F	114—124	**0.602**	0.411	0.084	0.472	0.442
SS-bonds						
SS-bonds fragments	26—84	**0.534**	**0.551**	0.308	0.378	0.836
	40—95	**0.633**	**0.611**	0.205	0.312	0.807
	58—110	**0.536**	**0.565**	0.304	0.431	0.848
	65—72	0.437	**0.563**	0.165	0.473	0.916

Table 6.A.2 Values of fuzzy oil drop parameters calculated for the structure of ribonuclease (5RSA) and its selected fragments. Values printed in boldface reflect discordance expressed either by high RD values or biased domination of HvO correlation coefficient. The leftmost column lists secondary folds and (where applicable) the location of catalytic residues. Underscored values are regarded as particularly important for determination of conformational characteristics.—cont'd

Secondary structure						
β-strands	42—48	0.443	0.381	0.200	0.512	0.572
	79—87	**0.542**	0.530	0.389	0.457	0.878
	96—105	**0.548**	0.359	**0.087**	**-0.013**	**0.862**
β-sheet		**0.557**	0.443	**0.386**	**0.378**	**0.811**
β-strands	61—63	0.342	0.163	0.870	0.947	0.981
	71—75	0.459	**0.863**	0.289	0.349	0.980
	106—111	**0.575**	0.487	0.568	0.168	0.792
119H, 120F	**116—124**	**0.657**	0.280	-0.093	0.228	0.276
β-sheet		**0.591**	**0.530**	0.337	0.227	0.738
Helices 12H	**3—13**	**0.568**	0.477	**0.249**	**0.304**	**0.887**
	24—33	0.320	0.503	0.763	0.707	0.939
	50—56	0.477	0.272	0.514	0.706	0.891
	57—60	0.201	0.412	0.758	0.828	0.945

in the case of the lysozyme, this micellar "capsule" protects the active site, which—in all likelihood—must remain discordant to ensure a suitable environment for catalysis.

In summary, we may note that each of the presented enzymes includes fragments which resemble an organized micelle. In these fragments the observed distribution of hydrophobicity is consistent with the theoretical Gaussian, and they may be regarded as a spherical (or globular) micelle.

In contrast, the fragments marked in red exhibit major deviations from the theoretical distribution of hydrophobicity. By actively opposing micellization, such fragments carry information which the protein requires to fulfill its biological purpose.

Unlike surfactant micelles which consist of identical unit molecules and retain perfect symmetry, the protein may be described as an "intelligent micelle". Local deviations from symmetric patterns play an important role in this context.

FOD-based folding simulations carried out for these two proteins are described in Refs. [3,4].

References

[1] Artymiuk PJ, Blake CCF. Refinement of human lysozyme at 1.5 Å resolution analysis of non-bonded and hydrogen-bond interactions. Journal of Molecular Biology 1981; 152(4):737–62. https://doi.org/10.1016/0022-2836(81)90125-x.

[2] Wlodawer A, Borkakoti N, Moss DS, Howlin B. Comparison of two independently refined models of ribonuclease-A. Acta Crystallographica Section B Structural Science 1986;42(4):379–87. https://doi.org/10.1107/s0108768186098063.

[3] Jurkowski W, Bryliński M, Konieczny L, Roterman I. Lysozyme Folded In Silico According to the limited conformational sub-space. Journal of Biomolecular Structure and Dynamics 2004;22(2):149–57. https://doi.org/10.1080/07391102.2004.10506991.

[4] Jurkowski W, Brylinski M, Konieczny L, Wiśniowski Z, Roterman I. Conformational subspace in simulation of early-stage protein folding. Proteins: Structure, Function, and Bioinformatics 2004;55(1):115–27. https://doi.org/10.1002/prot.20002.