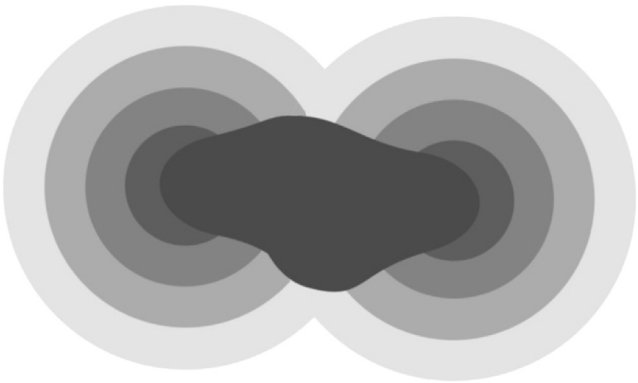
CHAPTER 6.B

Protein-protein interaction encoded as an exposure of hydrophobic residues on the surface

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



According to the fuzzy oil drop model, information required for complexation of an external polypeptide chain should be encoded as a local exposure of hydrophobicity on the protein surface. Of course, this is not the sole mechanism by which protein complexes may emerge; however, exposed hydrophobicity appears to play an important part in the complexation process. A local excess of hydrophobicity — especially if

From Globular Proteins to Amyloids
ISBN: 978-0-08-102981-7
https://doi.org/10.1016/B978-0-08-102981-7.00008-7

occurring on the surface — creates an environment which favors interaction with other polypeptide chains; particularly those, which also expose a hydrophobic interface.

This chapter discusses a homodimer which meets all the above conditions: DNA-binding regulatory protein — *Escherichia coli* met repressor in complex with a co-repressor (S-adenosylmethionine) and DNA (PDB ID: 1CMA) [1], as well as in the absence of these molecules (PDB ID: 1CMB) [2], i.e. in the apo form.

To reveal the role of the complexation interface we calculate model parameters for the complex, for individual chains treated as components of the complex and for standalone individual chains (adjusting the shape of the 3D Gaussian capsule as necessary). In addition, we determine the status of interface residues in the complex as well as in individual chains.

As shown in Table 6.B.1, the apo dimer is consistent with the model, while the corresponding structure, which includes the co-repressor and the DNA fragment, exhibits deviations from the model. A common property of the apo dimer and the larger complex is that their interfaces are FOD-accordant (in the area of inter-chain interactions) while the remainder of the structure exhibits elevated RD values, which indicates that the interface plays an important role in stabilizing the complex as a whole. When deprived of interface residues, the dimers exhibit poor structural stability.

As a whole, the apo dimer is accordant, while the co-repressor complex is not. When attempting to identify the causes of this discrepancy, we should note the fragment at 88—99, which is locally discordant (i.e. eliminating it from calculation significantly lowers the RD value for the remaining part of the chain).

When treated as a standalone structure, each monomer exhibits characteristic local deviations between O and T. In all cases these deviations may be attributed to residues which comprise the interface. Clearly, the structure of the interface is entropically disadvantageous in a monomer — indeed, its purpose is to encode information which would facilitate the formation of a dimer.

The above hypothesis may be criticized by noting that we are analyzing each chain as it appears in the dimer, and that, when analyzed on its own, it might not exhibit such an anomalous distribution of hydrophobicity. This can be countered by noting that even when chains undergo structural realignment during dimerization, they must first possess the capacity for such realignment, and that the information required in this process must be encoded in their structure.

Table 6.B.1 Status of the apo dimer (1CMB) and the dimer in complex with a co-repressor (1CMA), computed for the whole dimer, for individual chains treated as part of the dimer and for standalone individual chains respectively. The table also lists the corresponding values which result from elimination of interface residues (no P—P) and for the interface residues themselves (P—P), as well as elimination of other selected fragments (to identify the micellar part of the protein). Values listed in boldface indicate discordance.

Complex in apo form (1CMB)

		RD		Correlation coefficient		
Apo complex (1CMB)	**Fragment**	**T-O-R**	**T-O—H**	**HvT**	**TvO**	**HvO**
Complex		0.489	0.402	0.371	0.616	0.766
	Chain A	0.449	0.351	0.378	0.649	0.757
	Chain B	**0.522**	0.446	0.365	0.585	0.776
Complex						
No P—P		**0.554**	0.450	0.108	0.412	0.734
P—P		0.380	0.214	0.465	0.700	0.728
Chain A						
No P—P		**0.531**	0.406	0.038	0.400	0.712
P—P		0.336	0.173	0.520	0.760	0.718
Chain B						
No P—P		**0.570**	0.487	0.170	0.427	0.753
P—P		0.419	0.257	0.401	0.630	0.738

(*Continued*)

Table 6.B.1 Status of the apo dimer (1CMB) and the dimer in complex with a co-repressor (1CMA), computed for the whole dimer, for individual chains treated as part of the dimer and for standalone individual chains respectively. The table also lists the corresponding values which result from elimination of interface residues (no P—P) and for the interface residues themselves (P—P), as well as elimination of other selected fragments (to identify the micellar part of the protein). Values listed in boldface indicate discordance.—cont'd

Complex in apo form (1CMB)

		RD		Correlation coefficient		
Apo complex (1CMB)	**Fragment**	**T-O-R**	**T-O—H**	**HvT**	**TvO**	**HvO**
Chain A — individual						
Chain A		**0.568**	0.453	0.148	0.415	0.720
No P—P		**0.530**	0.405	0.073	0.352	0.714
P—P		**0.658**	**0.561**	0.265	0.558	0.698
DNA binding		0.473	0.182	0.174	0.762	0.625
No DNA bin.		**0.571**	0.479	0.152	0.395	0.708
Eliminated residues	45-47, 95-100, 68—72, 79—83	0.453	0.297	0.221	0.441	0.711
Chain B — individual						
Chain B		**0.586**	0.488	0.137	0.367	0.730
No P—P		**0.539**	0.450	0.134	0.356	0.750
P—P		**0.701**	**0.603**	0.134	0.417	0.692

Complex including co-repressor (1CMA)						
1CMA	Fragment	RD		Correlation coefficient		
Complex incl. co-repressor		T-O-R	T-O−H	HvT	TvO	HvO
Complex		**0.518**	0.428	0.344	0.570	0.763
	Chain A	**0.520**	0.437	0.346	0.560	0.757
	Chain B	**0.516**	0.420	0.343	0.581	0.770
Eliminated residues	88−99	0.388	0.310	0.400	0.689	0.750
Chain A + Chain B − status in complex						
No P−P		**0.578**	0.447	0.164	0.410	0.726
P−P		0.423	0.318	0.441	0.615	0.770
Chain A		**0.520**	0.437	0.346	0.560	0.757
No P−P		**0.578**	0.454	0.144	0.399	0.714
P−P		0.436	0.333	0.450	0.590	0.763
Chain B		**0.516**	0.420	0.343	0.581	0.770
No P−P		**0.578**	0.440	0.183	0.421	0.739
P−P		0.410	0.302	0.432	0.641	0.778
Chain A − individual						
Chain A		**0.592**	0.485	0.149	0.359	0.716
No P−P		**0.557**	0.432	0.178	0.349	0.713
P−P		**0.688**	**0.671**	0.143	0.440	0.709
Eliminated residues	P−P and 3−5,88-99	0.481	0.311	0.260	0.555	0.657

(*Continued*)

Table 6.B.1 Status of the apo dimer (1CMB) and the dimer in complex with a co-repressor (1CMA), computed for the whole dimer, for individual chains treated as part of the dimer and for standalone individual chains respectively. The table also lists the corresponding values which result from elimination of interface residues (no P—P) and for the interface residues themselves (P—P), as well as elimination of other selected fragments (to identify the micellar part of the protein). Values listed in boldface indicate discordance.—cont'd

Complex in apo form (1CMB)						
		RD		Correlation coefficient		
Apo complex (1CMB)	**Fragment**	**T-O-R**	**T-O—H**	**HvT**	**TvO**	**HvO**
Chain B — individual						
Chain B		**0.587**	**0.645**	0.134	0.356	0.728
No P—P		**0.561**	0.421	0.195	0.333	0.736
P—P		**0.658**	**0.619**	0.040	0.476	0.683
Eliminated residues	P—P and 3—5, 33, 88-99	0.476	0.294	0.318	0.557	0.687

It is reasonable to conclude that a hydrophobic area — especially when located on the molecular surface — may mediate complexation with external proteins. Of course, this is not the only means of identifying complexation interfaces. The CAPRI project [3] shows that it is possible to pinpoint interfaces by studying nonbonding interactions and geometric properties of monomers; nevertheless, the presented line of reasoning, based on the fuzzy oil drop model, may provide useful results at least for a subset of protein complexes.

A detailed discussion of how various types of homodimers are formed can be found in Ref. [4], which distinguishes several types of relations between monomers depending on FOD accordance/discordance of the resulting dimer. This study proposes a new dimerization mechanism which is not strictly based on geometric alignment, but instead relies on the presence of a "quasi-domain" in the interface area, to which both monomers contribute, and which remains highly consistent with the fuzzy oil drop model. This domain is believed to structurally stabilize proteins which must withstand external forces — for example dystrophin, which connects the cytoskeleton to muscular proteins and is therefore subject to frequent stretching.

Table 6.B.1 reveals that, in some cases, eliminating interface residues does not cause RD to become lower than 0.5. It might be interesting to speculate about other possible causes of this phenomenon. As it turns out, the fragment at 88–99 is particularly discordant in nearly all presented structures (especially when a co-repressor is involved). As shown in the 3D presentation, this discordance may be attributed to the exposed outlying loop, which is structurally flexible and may undergo frequent conformational changes.

Residues involved in dimerization, distinguished in Fig. 6.B.1, are typically characterized by excess hydrophobicity. This is particularly true of the fragment at 20–40. Other fragments which diverge from the theoretical distribution (marked in Fig. 6.B.1) are either too hydrophobic or insufficiently hydrophobic given their location. Eliminating them from calculations lowers the corresponding RD values (Figs. 6.B.2 and Fig. 6.B.3).

Using a homodimer as a representative example shows how local deviations from the theoretical distribution of hydrophobicity may assist dimerization. Given that the interface is highly accordant in the dimer, we may conclude that — in the scope of an individual chain — it represents a targeted form of discordance, specifically constructed to enable interactions with the intended complexation partner (Fig. 6.B.4)

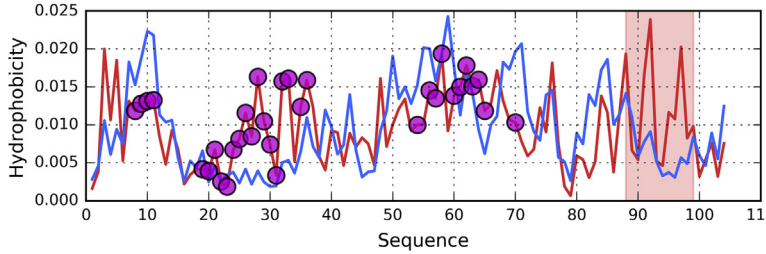


Fig. 6.B.1 Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for 1CMB chain A. Magenta circles mark residues involved in P-P interaction. Red background indicates a discordant fragment (88—99), which must be eliminated along with interacting residues to produce a value of RD below 0.5.

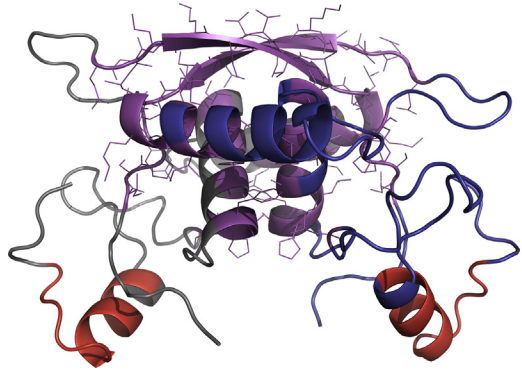


Fig. 6.B.2 3D presentation of 1CMB: chain (A) blue, chain (B) gray. Residues engaged in protein-protein interactions are magenta-colored and have side chains displayed. The fragments marked in each chain in red (88—99) are locally discordant.

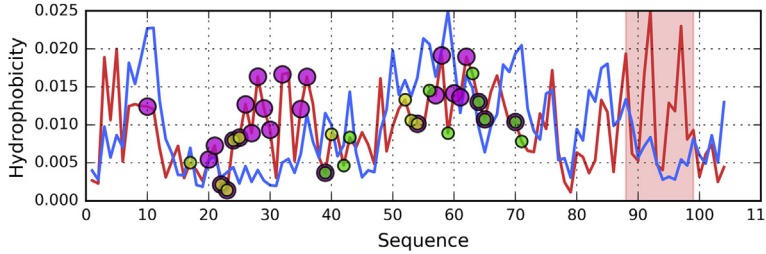


Fig. 6.B.3 Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for 1CMA chain A. Circles mark residues involved in interaction with other compounds in the complex: magenta — chain B (P—P), green — ligand, yellow — DNA. Red background indicates a discordant fragment (88—99), which when eliminated the RD value decreases below 0.5.

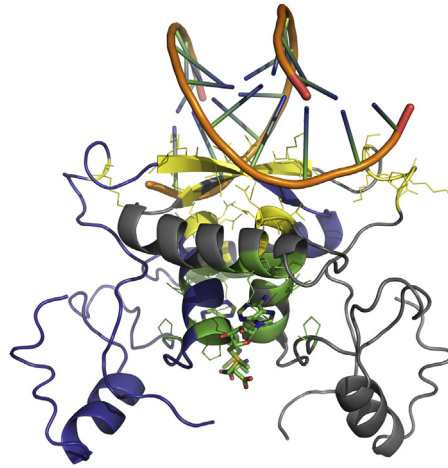


Fig. 6.B.4 3D presentation of 1CMA: chain (A) blue, chain (B) gray. Residues engaged in protein-ligand (green) and protein-nucleic (yellow) interactions have side chains displayed. The fragments marked in each chain in red (88–99) are locally discordant.

When considering the dimer as a whole, each chain is also regarded as accordant, which shows that both chains cooperate to build a shared hydrophobic core. Again, the information required in the process of complexation is expressed as discrepancies between T and O in each individual chain. Note that dimerization is not the ultimate goal of the complex, and that it also includes residues which mediate its biological activity – i.e. DNA binding.

The above conclusions remain valid both for the apo form of the presented protein and for its complex with the co-repressor, suggesting that the FOD status is related to the biological role of the complex. A discussion of various types of relations between the monomer and the dimer in light of the fuzzy oil drop model can be found in Ref. [4].

The status of the homodimer suggests that this structure emerges as a result of interactions with the water environment, which produces a shared hydrophobic core. When analyzed as part of the dimer, chain A appears to contribute to the shared core, while chain B is discordant on its own and does not exhibit accordance when viewed as part of the complex. Balanced values of correlation coefficients (TvO and HvO) indicate a synergistically optimized structure. It is also interesting to note the very low RD value computed for the interface, showing that the interface area is a major stabilizing factor, enabling the dimer to retain a shared core. Eliminating interface residues from FOD calculations produces a much more discordant dimer

structure, with a notable increase in the value of RD. It follows that the dimer functions as a coherent structural unit, stabilized by its own hydrophobic core to which both chains contribute.

Individual chains

Each chain, when analyzed on its own, appears to diverge from the monocentric distribution of hydrophobicity. This is linked to exposure of strongly hydrophobic residues on the surface, contrary to the FOD model. Notably, the exposed residues belong to the complexation interface and are expected to encode information required in the complexation process. This conclusion is supported by calculation of RD values for interface fragments: eliminating such residues from computations results in a significant decrease in RD, although the remainder of each chain is still regarded as discordant.

The activity profile of the presented protein involves DNA binding. Residues which mediate this process conform to the model since they exhibit high polarity and are exposed on the surface, in agreement with the micellar protein structure model. Eliminating these residues results in a higher value of RD for chain A (when treated as a separate unit).

DNA binding

Identification of residues which cause RD to exceed 0.5 despite prior elimination of discordant interface residues may yield clues regarding the protein's capability for structural rearrangement. In the scope of the individual chain, discordance is caused by the fragments at 45–47, 68–72, 79–83 and 95–100 (in addition to the aforementioned interface fragments). All these fragments belong to outlying loops and may potentially undergo conformational changes during DNA complexation.

The sketch included in the chapter title illustrates the specific conditions discussed above. A detailed analysis of discrepancies between the status of individual monomers and the resulting dimers (in all possible variants) is presented in Refs. [4–7].

References

[1] Somers WS, Phillips SEV. Crystal structure of the met represser–operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. Nature 1992;359(6394): 387–93. https://doi.org/10.1038/359387a0.

[2] Rafferty JB, Somers WS, Saint-Girons I, Phillips SEV. Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. Nature 1989; 341(6244):705–10. https://doi.org/10.1038/341705a0.

[3] Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJE, Vajda S, Wodak SJ. CAPRI: a critical assessment of PRedicted interactions. Proteins: Structure, Function, and Genetics 2003;52(1):2–9. https://doi.org/10.1002/prot.10381.
[4] Dygut J, Kalinowska B, Banach M, Piwowar M, Konieczny L, Roterman I. Structural interface forms and their involvement in stabilization of multidomain proteins or protein complexes. International Journal of Molecular Sciences 2016;17(10):1741. https://doi.org/10.3390/ijms17101741.
[5] Marchewka D, Jurkowski W, Banach M, Roterman-Konieczna I. Prediction of protein-protein binding interfaces. Focus on Structural Biology 2012:105–33. https://doi.org/10.1007/978-94-007-5285-6_6.
[6] Banach M, Konieczny L, Roterman-Konieczna I. Can the structure of the hydrophobic core determine the complexation site? Focus on Structural Biology 2012:41–54. https://doi.org/10.1007/978-94-007-5285-6_3.
[7] Banach M, Konieczny L, Roterman-Konieczna I. Use of the "fuzzy oil drop" model to identify the complexation area in protein homodimers. Protein Folding in Silico 2012: 95–122. https://doi.org/10.1533/9781908818256.95.