Protein–DNA and protein–protein interactions in λ-repressor/operator system: A review

Siddhartha Roy

Department of Biophysics, Bose Institute, P 1/12 C.I.T. Scheme VII M, Calcutta 700 054, India.

λ-repressor is a two-domain protein. The N-terminal domain interacts with the DNA, while the C-terminal domain is responsible for all protein–protein interactions. Functioning of the repressor in the natural context implied that the two domains should have flexibility. Denaturation and fluorescence anisotropy studies, however, demonstrate that the situation is more complex. It is likely that in the free repressor the two domains interact with each other and are not completely free to rotate. The binding of operator to N-terminal domain, however, causes a conformational change in the C-terminal domain. This conformational change leads to change in protein–protein interaction. One of the major changes that is seen, is that of the role of the C-terminal tail region in protein–protein interaction. It is hypothesized that the role of the conformational change is to couple the specific operator binding information to the protein–protein interaction, which may lead to increased specificity at the required sites.

REGULATION of information flow from DNA to RNA to protein is at the center of all biological phenomena. This regulation, both when following an internal programme and in response to an external stimulus, is primarily at the level of transcription. The regulation of transcription is a very complex process in eukaryotes and is not well understood. Even the basal transcription in yeast may take tens of protein factors[1]. Such complex systems are difficult to study from structural points of view with the present level of technology. Prokaryotic and phage regulatory systems are often much simpler and at the same time incorporate many of the regulatory features of the higher eukaryotes[2,3]. Thus, they are ideally suited for investigations by structural and biophysical techniques.

Many of the prokaryotic regulatory systems have been studied in fair details by genetic and molecular biological techniques. They include *Lac, Gal, Trp, deo* and *Ara* operon of *E. coli*[7] and lytic-lysogenic regulatory system of bacteriophage λ[8–10]. Protein components of these and several other prokaryotic regulatory systems are well characterized and nature of the regulatory events have been described. Crystal and NMR structures of some of these proteins and some fragments have also been elucidated[11–14]. Thus, several of these systems are poised for further investigations of the regulatory events in quantitative terms and at atomic levels.

My laboratory has focused attention on the lytic-lysogenic switching system of bacteriophage λ. Extensive genetic and biochemical work from the early '70s by Mark Ptashne and coworkers has established the components and the basic nature of the system. The switching system consists of three contiguous 17 base pair sequences known as the right operator sites, O_R1, O_R2 and O_R3, and two proteins, λ-repressor and *cro*. This collection of operator sites, known as the right operator or O_R, also contains two promoters, P_{RM} and P_R (Figure 1). The structural gene of the λ-repressor is under the control of the promoter P_{RM} and the structural gene of the *cro* is under the control of the promoter P_R.

The cro protein is a homodimer of a polypeptide chain of 66 amino acids and binds to the three operator sites, in the order of decreasing affinity, $O_R3 > O_R2 > O_R1$ (ref. 15). The cro protein functions as a down regulator of transcription[16]. In contrast, the λ-repressor is a multifunctional protein that acts as both *up* and *down* regulator depending on the context[17,18]. In addition, λ-repressor binds to the three operator sites, co-operatively, in the order of decreasing affinity $O_R1 > O_R2 > O_R3$ (ref. 19). This co-operative binding is vital for the functioning of the genetic switch, since, otherwise normal non-cooperative mutants lose their ability to lysogenize completely. In the lysogenic state, two repressor molecules occupy O_R1 and O_R2, and activate the promoter P_{RM} and repressing the promoter P_R, thus shutting off *cro* synthesis and continuing λ-repressor synthesis. The continuing repressor synthesis is required to replenish the dilution losses due to cell division and growth. No other gene is transcribed in this state and the steady-state is maintained. Following DNA damage, the recA protein cleaves the repressor, destroying the co-operative binding and weakening the affinity for the operator sites. This throws the switch, shutting off the repressor synthesis and initiating *cro* synthesis, resulting in entry into the lytic pathway (Figure 2).

It is clear from the above description of the operation of the switch, that the λ-repressor plays the central role in switching the phage from lysogenic to the lytic mode

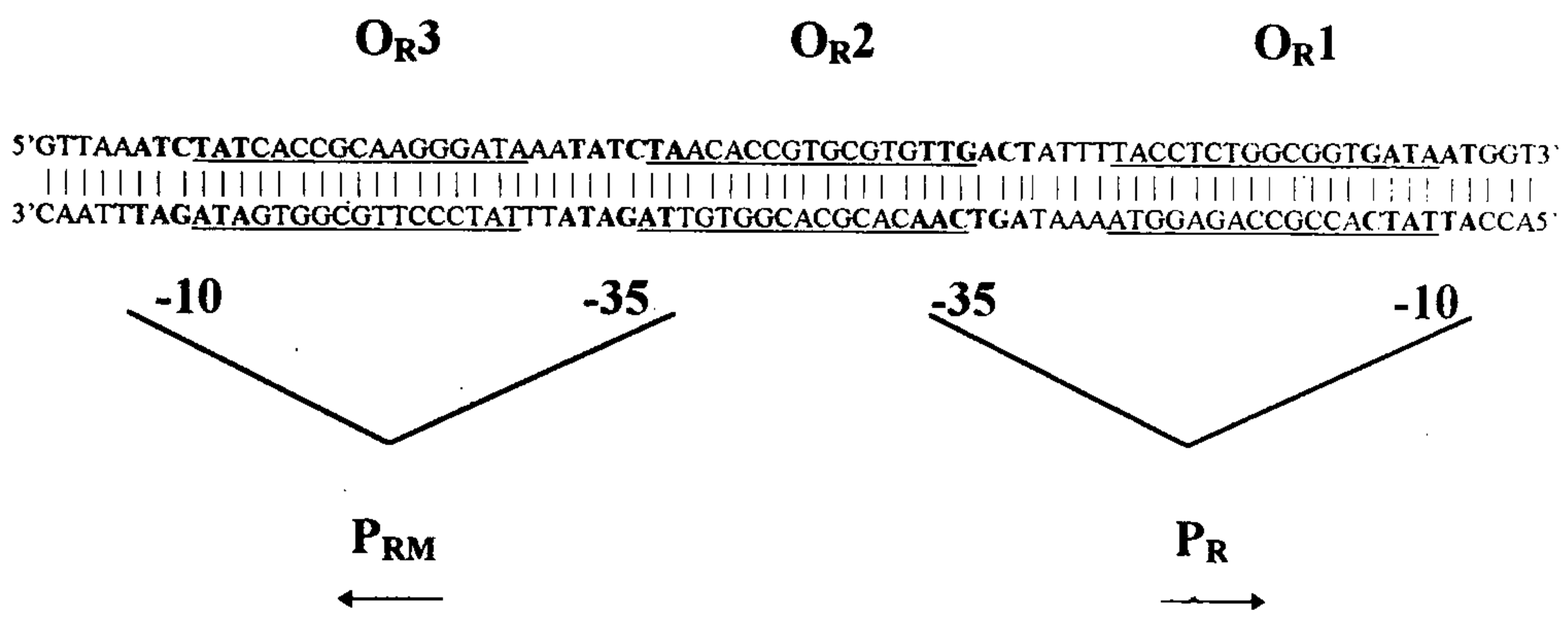


Figure 1. DNA sequence of right operator region of bacteriophage λ.

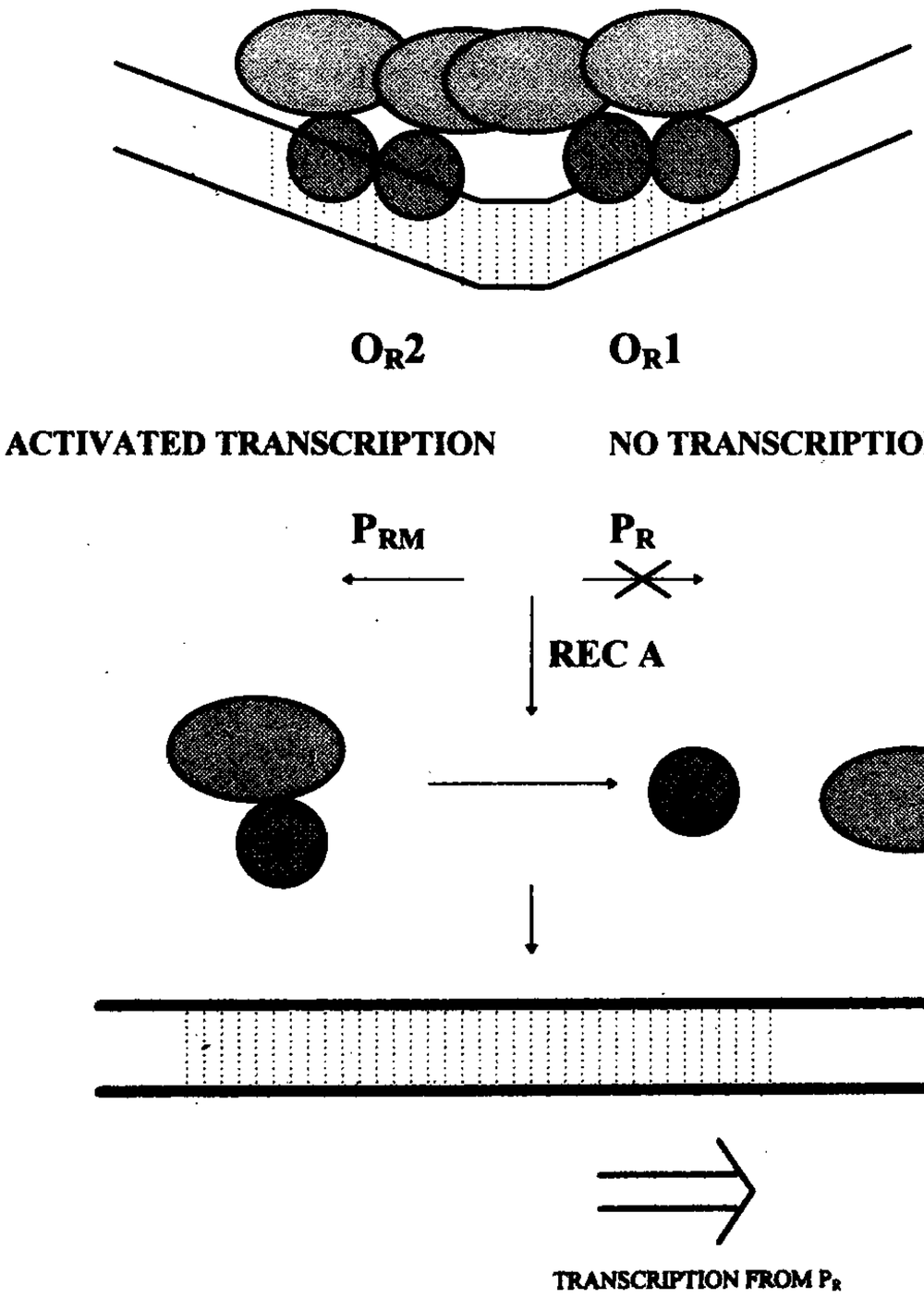


Figure 2. Depiction of O_R1–O_R2/λ–repressor complex in lysogenic state and subsequent induction to lytic pathway. The upper half of the figure represents two repressor molecules bound to the operator sites O_R1–O_R2 and preventing transcription from P_R but activating transcription from P_{RM}. The middle part of the figure depicts the cleavage of the repressor by RecA and the bottom part represents the unbound form of DNA in which transcription from P_R takes place.

of growth. λ-repressor monomer consists of a single polypeptide chain of 236 amino acids. The monomers associate to form dimers, tetramers, octamers and at very high protein concentrations, higher order structures[20,21]. The chain folds into two domains, 1–92 and 132–236 with a hinge region in between the two domains[20]. The N-terminal domain contains the helix–turn–helix motif, commonly found in many prokaryotic DNA-binding proteins, and has been shown to bind operator DNA[22]. The C-terminal domain is largely responsible for dimer and higher order aggregate formation as well as protein–protein contact needed for co-operative binding to multiple operator sites[23]. Very little is known about the structure of the hinge region or its function.

A problem of great magnitude, such as regulation of gene expression and genetic switching, requires study of many systems. Understanding of their diversity and underlying unity is essential to unravel the complexity of the problem. One or a few emerge as the paradigm. *Lac* and λ regulatory systems have perhaps emerged as the best characterized systems in terms of genetics, biochemistry and structure. Thus, the stage was set for quantitative biophysical studies to understand the regulation in quantitative terms. Seminal work by Gary Ackers and co-workers[24–26] in the eighties has established the basic physicochemical parameters of the system, such as binding affinities, cooperative interaction energies, etc. The task now is to elucidate conformational and dynamical aspects of the system, so that finally it can be understood at the residue or the atomic level.

My laboratory has undertaken a long term programme to investigate the biophysical aspects of functioning of the λ genetic switch, in general and the interactions of λ-repressor, the most important component of the switch, in particular. The interactions of λ-repressor with the operator sites have been studied in detail by Gary Ackers and his co-workers, using quantitative footprinting[24,25]. They measured the net co-operative interaction energies, which turn out to be small, around 2–3 kcal/mole. This net co-operative interaction energy consists of at least the following elements: (i) The

proximity effect of bringing the two adjacent molecules to one another. This effect, in analogy to the chelate effect, should be entropic in origin. (ii) The protein–protein interaction energy while bound to the two operator sites. (iii) The DNA binding energy. (iv) The protein distortion energy. The last two effects oppose the first two. The decomposition of the net co-operative interaction energy to its constituent parts is essential if one has to understand co-operatively in molecular and atomic terms. Thus, we have focused our attention on these individual effects, concentrating on protein–protein interaction, DNA binding and DNA-induced conformational change.

Domain–domain interaction in λ-repressor

It was generally believed that the two domains of the repressor may be free to rotate, with the hinge acting as a connecting flexible region. This putative flexibility was thought to be important in the functioning of the repressor since, the repressor can co-operatively bind to two operator sites separated by variable number of base pairs. The crystal structure of the isolated N-terminal domain as well as its complex with the O_L1, is known[12,22]. Very little information on the C-terminal domain and hinge structure is, however, available. NMR studies of the whole repressor molecules suggested motional freedom of the N-terminal domain due to the observed line narrowing[27]. The experiments were, however, done at very high protein concentrations (3–5 mM) where λ-repressor is known to oligomerize to higher order structures, which may have properties different from those of the free dimer. Equilibrium denaturation of λ-repressor by urea has been studied using fluorescence and circular dichroism spectroscopy[28]. λ-repressor contains three tryptophan residues, 129, 142, 230 and three cysteine residues, 180, 215 and 219. The λ-repressor denatures in three distinct phases: in the first phase, which is centered around 2.0 M, there is a significant red shift of tryptophan fluorescence, but little change in far-UV circular dichroism spectrum. The second phase, centered around 3.5 M urea, is characterized by large decrease of the CD spectral intensity, but no change in the tryptophan fluorescence spectrum. The third phase occurs at very high urea concentrations, above 8 M urea. Fluorescence polarization studies indicate that the repressor remains as a dimer and the dissociation into monomers occurs around 7.5 M urea. As mentioned above the λ-repressor contains three cysteine residues, all of which are situated in the C-terminal domain. The attempted reaction of the cysteines with DTNB indicated that the cysteines were unreactive in the native state. Exhaustive protease digestion, however, produced near-full reactivity, suggesting inaccessibility as the cause of inertness. This lack of reactivity is preserved even at 6 M urea, suggesting that the first two phases of denaturation do not involve total denaturation of the C-terminal domain. One sulfhydryl group becomes available for reaction at urea concentrations higher than 8 M. The intact repressor can be cleaved by papain into 1–92 and 132–236 fragments, which can be isolated[20]. Denaturation of the isolated N-terminal domain coincides with the second phase of the denaturation, suggesting that the second phase is the denaturation of the N-terminal domain. This also agrees with the lack of tryptophan residues in this domain (no change in tryptophan fluorescence) and high alpha helical content (large change in CD signal). The nature of the first transition thus could be partial denaturation of the C-terminal domain or the hinge region or both and hence, denaturation of the isolated 132–236 fragment was studied. This fragment still shows the first phase of the denaturation, only the mid-point shifts to around 1.0 M urea concentration. This suggests that the first phase of the denaturation involves denaturation of the part of the C-terminal domain. Under controlled proteolysis condition, the repressor can be cleaved by papain to produce the fragment 93–236 (ref. 20). This fragment has also been purified. The equilibrium denaturation shows that the first denaturation phase is centered around 1.4 M urea. These data suggest that the presence of the hinge region and the N-terminal domain confers additional stability to the part of the C-terminal domain which undergoes denaturation in low urea concentrations. Since the fragment 132–236 preserves the first transition involving tryptophan residues, W142 or W230 must be responsible for the red shift. N-Bromosuccinimide modification studies suggest that W230 is highly red shifted (λ_{max} around 345 nm). Since the first transition involves red shift of one tryptophan residue having λ_{max} of less than 340 nm, it suggests that W142 may be the tryptophan involved in the first transition. Thus, it may be concluded that the N-terminal domain, the hinge and the N-terminal part of the C-terminal domain may interact with each other in the native repressor and the two domains may not be completely isolated. The interaction energy, however, may be small (approximately 1 kcal/mole; obtained from the extrapolation to zero urea concentrations), suggesting that the freedom of internal rotation may be gained by sacrifice of only around 1 kcal/mole energy.

The concept of interacting domains enunciated above received strong backing from fluorescence polarization studies. As mentioned above, the repressor lacked reactive cysteine groups to be tagged with appropriate fluorophores. The protein may be labelled with dansyl chloride without the loss of biological activity[29]. Rotational correlation time, under conditions known to favour the dimer, was measured from Perrin plot. The result obtained was consistent with a globular dimeric protein of molecular weight of 52 kDa. If the domains

were completely free to rotate with respect to each other, then this would have led to smaller values of the rotational correlation time.

Self-assembly of λ-repressor

The preparation of dansyl labelled repressor with full biological activity produced a unique opportunity to study protein association in λ-repressor. λ-repressor is known to be in a monomeric state at concentrations below nanomolar level. It dimerizes with an association constant of about 10^8 M^{-1} and remains predominantly dimeric till micromolar concentrations[26,30]. Early qualitative work suggested that beyond 1 μM, λ-repressor formed tetramer, octamer and eventually higher order structure[21]. The self-assembly was not known in quantitative terms. A quantitative measurement of the self-assembly is important to elucidate the nature of biologically important protein–protein interaction.

Using fluorescence anisotropy measurement, the assembly profile of λ-repressor from dimer to tetramer was determined. Under the solution conditions used (i.e. in the absence of divalent cations), the tetramer to octamer transitions was largely decoupled from the dimer to tetramer transition. The temperature dependence of the assembly profile showed the dimer to tetramer assembly to be highly temperature dependent and enthalpy driven with a ΔH of around –34 kcal/mole. The binding of the single operator site fragment O_R1 caused weakening of protein–protein interaction around ambient temperature. At about the same time other laboratories had reported a concerted transition from dimer to octamer approximately the same concentration range, using ultracentrifugation[31]. Their solution conditions, however, included millimolar concentrations of Ca^{++} and Mg^{++}. Controversy about the correct model of self-assembly lingered on for some more years, until recently, when two studies from different laboratories resolved the issue. Gary Ackers and co-workers have shown that the original concerted dimer to octamer model was based on inaccurate analysis of the ultracentrifuge data and the revised data showed much lower degree of coupling between the two transitions under the conditions previously reported to favour concerted dimer to octamer assembly (G. A. Ackers personal communication). Studies in our laboratory suggested that millimolar divalent cation concentrations promote significant coupling of the two transitions and in the absence of divalent cations only a small degree of coupling is seen[32]. Whether the effects of divalent cations on protein association have corresponding effect on protein–protein interaction responsible for binding co-operativity, is unknown at the present moment.

Operator-induced conformational change

The arguments presented above suggest that the N- and the C-terminal domains in λ-repressor may not be free to rotate. That raises an apparent dilemma. The various operator site pairs are separated by a variable number of base pairs (four through seven) and the repressor binds to these pairs of operator sites with co-operative interaction energies of similar magnitude. When the operator sites were separated by a much longer stretch of DNA on an artificial construct, the repressor still bound to these operator sites co-operatively[33]. Such insensitivity to the length of the intervening DNA suggests considerable flexibility in the C-terminal domain, since the N-terminal domain remains anchored to the operator site. It is possible that co-operative contact occurs between two operator sites bound dimers at the cost of N-terminal domain–C-terminal domain interaction. The other possibility is that binding of repressor to the operator site may lead to a global conformational change, leading to facilitated freedom of movement of the C-terminal domain. The possibility that binding of the operator site to the λ-repressor produces a global conformational change was investigated, using intrinsic tryptophan fluorescence and an environment-sensitive fluorescence probe, bis-ANS[34]. All the tryptophans of λ-repressor are situated away from the N-terminal domain and hence the DNA binding site. The binding of an oligonucleotide containing an isolated operator site, O_R1, leads to significant quenching of tryptophan fluorescence and a modest shift of emission maximum. Since all the tryptophans are situated away from the DNA binding site, the operator-induced change is likely to be transmitted to the C-terminal domain. Bis-ANS is an environment-sensitive probe with affinity for the apolar sites of proteins. Bis-ANS binds to the C-terminal domain of λ-repressor. Binding of the oligonucleotide containing O_R1 operator site causes fluorescence enhancement of the bound bis-ANS, indicating a global conformational change occurring upon binding of the oligonucleotide containing the operator site. Whether such a global conformational change leads to increased freedom of movement of the C-terminal domain, remains to be seen.

Role of C-terminal tail in protein association

During studies of protein association, it was observed that a shift of emission maximum of tryptophan fluorescence takes place during protein association at the concentration range that is known to promote dimer to tetramer association under the conditions used. This suggested that one of the three tryptophans may be at or near the protein–protein contact site and its identification may lead to localization of part of the dimer–dimer

contact site. Acrylamide quenching experiments showed that accessibility of at least one tryptophan is significantly different in the dimeric and tetrameric state[35]. This tryptophan also shifts its emission maximum from very red, 345 nm in the dimer to very blue 334 nm in the tetramer. N-Bromosuccinimide oxidation leaves one of the tryptophans unmodified in the tetrameric state, but not in the dimeric state. Peptide mapping, amino acid analysis and sequencing suggest that this tryptophan is W230. Based on this information we created a site-directed mutant F235C for attachment of fluorescence probe to the tail region of the protein. This sulfhydryl residue can be selectively labelled with an environment-sensitive probe, acrylodan. Acrylodan 235C-labelled protein showed significant shift of emission maximum and fluorescence enhancement upon association from dimer to tetramer. Thus, we conclude that the whole C-terminal tail region of the protein may be involved in the dimer–dimer association.

Interestingly, many non-cooperative mutants have been isolated by several groups and some of them are mapped in the C-terminal tail region of the protein[23,36]. But most of the mutants are mapped in the other parts of the C-terminal domain. Several of the mutant proteins have been studied in respect to their association properties and DNA-binding abilities. The only well-characterized tail mutation is S228N. Although this was originally characterized as a non-cooperative mutant[10], recent studies have shown that it is defective in monomer–dimer association and not co-operativity defective as was originally suggested[37,38]. Sedimentation velocity ultracentrifugation studies have shown that S228N does not associate beyond the dimer stage, i.e. aggregation defective. This study pointed towards the fact that the role of the tail in co-operative contact and free protein association may be different. Since there are two excellent markers, W230 and C235, for the tail region, we decided to explore this interesting issue further.

A double operator site, O_R1–O_R2, containing oligonucleotide was synthesized. Binding of λ-repressor to this oligonucleotide leads to protein–protein contact, while the repressor dimers remain bound to the two adjacent operator sites. This complex was studied using acrylamide quenching as a tool[35]. The acrylamide quenching pattern of the O_R1–O_R2/repressor complex is very similar to the dimer, indicating no change of environment of W230. Acrylodan-labelled protein at cysteine 235 also shows different behaviour. The AC235-repressor shows dramatic quenching of AC fluorescence upon single operator titration and similar effect is seen with double operator titration. This is in contrast to the fluorescence enhancement seen in the free repressor undergoing tetramer association. The thermodynamics of the protein association is also dramatically different in free protein tetramer formation and co-operative contact. Whereas the former is strongly enthalpy driven, the latter is iso-enthalpic and entropy driven. The difference in thermodynamics and the role of the C-terminal tail may be due to spatial constraint imposed upon the protein–protein association by the proximity of the two adjacent operator sites and rigidity of the DNA, or it may be due to a change in nature of the protein–protein interface, induced by operator binding. The protein–protein association in single operator site bound dimer is, like the co-operative interaction energy, iso-enthalpic and entropy driven. The behaviour of W230 is also unlike the free tetramer association in the operator bound dimer association. We may conclude from these studies that free dimer association is significantly different from co-operative dimer–dimer contacts and this difference is a result of the global conformational change induced by the binding of the operator.

Binding of λ-repressor also induces a conformational change in the double operator fragment O_R1–O_R2 as seen by difference circular dichroism spectra. This distortion is not seen in a control oligonucleotide of the same composition where the two operators have been placed on different sides of the DNA by insertion of half turn of the DNA. This suggests that co-operative interaction of the λ-repressor is responsible for the distortion of the DNA and the distortion probably originates from the intervening stretch between O_R1 and O_R2. This result is consistent with DNAse I hypersensitive band seen in intervening sequence by quantitative footprint experiments[39].

Significance of the operator-induced conformational change

Based on the fact that the operator binding causes a global conformational change in the repressor, which ultimately leads to change of protein–protein interaction and DNA distortion induced by protein–protein co-operative contact, we propose that the λ-repressor/O_R1–O_R2 complex is a strongly coupled system, in which operator binding, DNA distortion and protein–protein interaction are strongly coupled. This can be understood in greater detail from the thermodynamic cycle depicted in Figure 3. The sum of intrinsic interaction energies (including entropic gain due to proximity effect) is not fully expressed, and the observed binding energy is less.

$$\Delta G_{ob} = \Delta G_{dis} + \Delta G_{bin} + \Delta G_{int}.$$

ΔG_{ob} is lower than $\Delta G_{int} + \Delta G_{bin}$. The difference is used to pay for the distortion of the DNA and protein. Clearly the central question is, what is the role of DNA and protein distortion in the system?

One possible answer is that the distortions are needed to bring the system in right geometric disposition. We feel that this is a red herring, and the real significance lies elsewhere. It was previously observed that

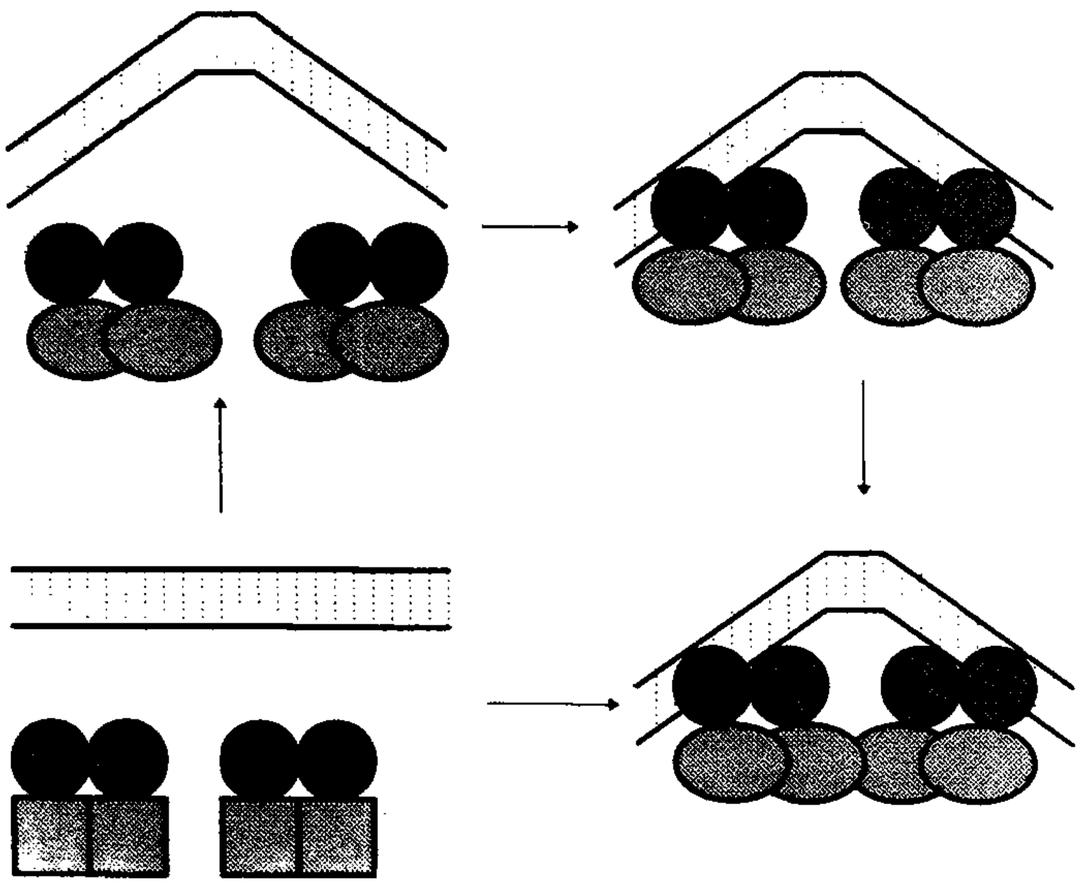


Figure 3. Decomposition of net co-operative interaction energy into individual interaction energies using binding cycle. The transformation from unbound DNA and unbound protein (lower left corner) to co-operatively bound complex (lower right hand corner) releases ΔG_{ob}. The free repressor C-terminal domain is represented by a square, whereas the operator-bound conformation is represented by an ellipse. The transformation from lower left to upper left, represents distortion of DNA and protein and is represented by ΔG_{dis}. The transformation from upper left to upper right represents binding energy of the distorted states, ΔG_{bin}. The transformation from upper right to lower right represents the protein–protein interaction energy while the repressor molecules are bound to the two adjacent operators, ΔG_{int}.

non-specific DNA sequences do not induce the conformational change induced by specific operator sequence. This suggests that the conformational change may be a mechanism to couple the correct recognition of sequence to appropriate protein–protein interaction and subsequent complex formation. This may prevent formation of a complex at inappropriate sequences.

Is there any other possible role for the conformational change and strong couple? In the classical model of *down regulation*, the repressor down regulates transcription by excluding the polymerase from the promoter. Recent work has suggested that this may not be the case[40]. A number of studies have shown that polymerase and repressor can coexist on the DNA[40,41]. Adhya and coworkers have hypothesized that like positive regulation, the negative regulation also takes place through repressor-polymerase protein–protein contact[2]. In a collaborative unpublished work[42] we have shown that under defined *in vitro* conditions, in the *gal* system, induction of transcription occurs at repressor concentrations much above the dissociation constant of the operator–repressor complex. This suggests that obligatory steric occlusion may not be the mechanism of *down regulation* in all cases and the repressor-polymerase contact should be considered as a serious possibility for the mechanism. If such a possibility turns out to be true, I speculate that another role for the operator-induced conformational change may be to present the correct interface for interaction with RNA polymerase, causing *down regulation*. In any non-specifically bound situation, such interface is not available for interaction with the RNA polymerase, thus decreasing the potential for causing harm elsewhere.

In summary, the regulatory system of bacteriophage λ has been investigated in detail from genetic, biochemical and to a moderate extent from structural points of view. Further structural investigations remain to be done. Quantitative biophysical parameters have been established from Gary Ackers' laboratory. We have shown the importance of conformational change and dynamics in the functioning of such a regulatory system. The challenge is now to understand the control and switching of transcription at the atomic level and in quantitative terms.

1. Ptashne, M., in *A Genetic Switch: Phage λ and Higher Organisms*, Cell Press, Boston, USA, 1992.
2. Adhya, S., *Annu. Rev. Genet.*, 1989, **23**, 227–250.
3. Schleif, R., *Annu. Rev. Biochem.*, 1992, **61**, 199–223.
4. Majumdar, A. and Adhya, S., *J. Biol. Chem.*, 1987, **262**, 13258–13262.
5. Dunn, T. M., Haber, S., Ogden, S. and Schleif, R. F., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 5017–5020.
6. Dandanell, G. and Hammer, K., *EMBO J.*, 1985, **4**, 3333–3338.
7. Eismann, E., von Wilcken-Bergmann, B. and Muller-Hill, B., *J. Mol. Biol.*, 1987, **195**, 949–952.
8. Sauer, R. T., Jordan, S. R. and Pabo, C. O., *Adv. Protein Chem.*, 1990, **40**, 1–61.
9. Hochschild, A. and Ptashne, M., *Cell*, 1986, **44**, 681–687.
10. Hochschild, A. and Ptashne, M., *Nature*, 1988, **336**, 353–357.
11. Friedman, A. M., Fischmann, T. O. and Steitz, T. A., *Science*, 1995, **268**, 1721–1727.

11a. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. and Lu, P., *Science*, 1966, **271**, 1247–1254.

12. Pabo, C. O. and Lewis, M. L., *Nature*, 1982, **298**, 443–447.
13. Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmonstein, R. Q., Luisi, B. T. and Sigler, P. B., *Nature*, 1988, **335**, 321–329.
14. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R. and Van Gunsteren, W. F., *J. Mol. Biol.*, 1985, **182**, 179–182.
15. Johnson, A., Meyer, B. J. and Ptashne, M., *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 1783–1787.
16. Takeda, Y., *J. Mol. Biol.*, 1979, **127**, 177–189.
17. Meyer, B. J. and Ptashne, M., *J. Mol. Biol.*, 1980, **139**, 195–205.
18. Meyer, B. J., Maurer, R. and Ptashne, M., *J. Mol. Biol.*, 1980, **139**, 163–194.
19. Johnson, A., Meyer, B. J. and Ptashne, M., *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 5061–5065.
20. Pabo, C. O., Sauer, R. T., Sturtvant, J. M. and Ptashne, M., *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 1608–1612.
21. Chadwick, P., Pirotta, V., Steinberg, R., Hopkins, N. and Ptashne, M., *Cold Spring Harb. Symp. Quant. Biol.*, 1970, **35**, 283–294.
22. Jordan, S. R. and Pabo, C. O., *Science*, 1988, **242**, 893–899.
23. Benson, N., Adams, C. and Youderian, P., *Mol. Microbiol.*, 1993, **11**, 567–579.
24. Senear, D. F., Brenowitz, M., Shea, M. A. and Ackers, G. K., *Biochemistry*, 1986, **25**, 7344–7354.

25. Brenowitz, M., Senear, D. F., Shea, M. A. and Ackers, G. K., *Proc. Natl. Acad. Sci. USA*, 1986a, **83**, 8462–8466.
26. Koblan, K. S. and Ackers, G. K., *Biochemistry*, 1991, **30**, 7817–7821.
27. Weiss, M. A., Karplus, M., Patel, D. J. and Sauer, R. T., *J. Biomol. Struct. Dyn.*, 1983, **1**, 151–157.
28. Banik, U., Saha, R., Mandal, N. C., Bhattacharyya, B. and Roy, S., *Eur. J. Biochem.*, 1992, **206**, 15–21.
29. Banik, U., Mandal, N. C., Bhattacharyya, B. and Roy, S., *J. Biol. Chem.*, 1993, **268**, 3938–3943.
30. Koblan, K. S. and Ackers, G. K., *Biochemistry*, 1991, **30**, 7822–7827.
31. Senear, D. F., Waxman, E., Laue, T. M., Eaton, S., Ross, J. B. A. and Rusinova, E., *Biochemistry*, 1993, **32**, 6179–6189.
32. Bandyopadhyay, S., Mukhopadhyay, C. and Roy, S., *Biochemistry*, 1996, **35**, 5033–5040.
33. Griffith, J., Hochschild, A. and Ptashne, M., *Nature*, 1986, **322**, 750–752.
34. Saha, R., Banik, U., Bandyopadhyay, S., Mandal, N. C., Bhattacharyya, B. and Roy, S., *J. Biol. Chem.*, 1992, **267**, 5862–5867.
35. Bandyopadhyay, S., Banik, U., Bhattacharyya, B., Mandal, N. C. and Roy, S., *Biochemistry*, 1995, **34**, 5090–5097.
36. Whipple, F. W., Kuldell, N. H., Cheatham, L. A. and Hochschild, A., *Genes and Development*, 1994, **8**, 1212–1223.
37. Burz, D. S. and Ackers, G. A., *Biochemistry*, 1994, **33**, 8406–8416.
38. Burz, D. S., Beckett, D., Benson, N. and Ackers, G. K., *Biochemistry*, 1994, **33**, 8399–8405.
39. Strahs, D. and Brenowitz, M., *J. Mol. Biol.*, 1994, **244**, 494–510.
40. Straney, S. B. and Crothers, D. M., *Cell*, 1987, **51**, 699–707.
41. Parrack, P. and Adhya, S., unpublished observation.
42. Chatterjee, S., Zhou, Y., Roy, S. and Adhya, S., unpublished observation.

Received 2 April 1996; revised accepted 5 June 1996

REVIEW ARTICLE

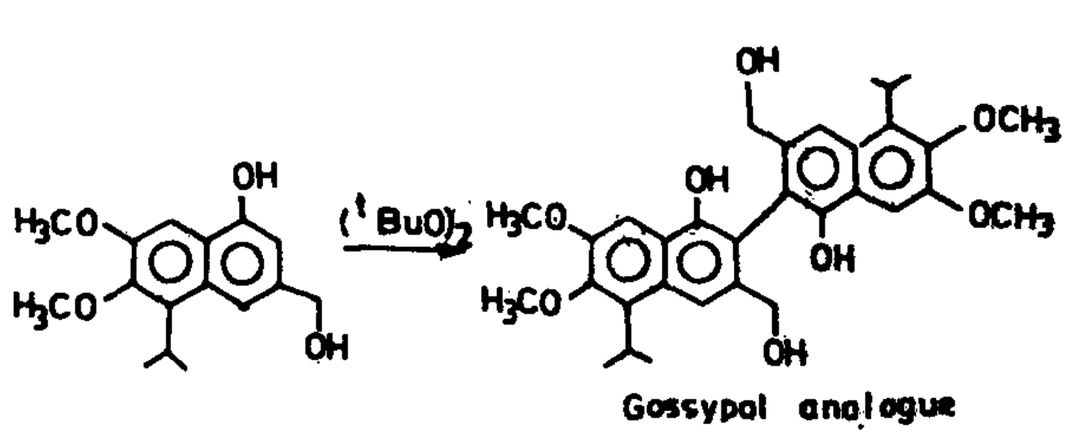
Oxidative phenol coupling: A key step for the biomimetic synthesis of many important natural products

Tarasankar Pal and Anjali Pal

Department of Chemistry, Indian Institute of Technology, Kharagpur 721 302, India

We present here the oxidative phenol coupling (OPC) reaction utilizing various reagents and leading to biomimetic synthesis of many natural polyphenolics.

OXIDATIVE phenol coupling (OPC) reaction incarnates a 'C–C' or 'C–O' bond between phenolic moieties through oxidation as delineated in all the schemes described here. OPC thus becomes a fundamental method for the synthesis of hydroxylated biaryls[1]. The reaction has previously been of little significance and synthetic value because it could not be controlled to the desired step. Unwanted side products were also formed if the aromatic substrate had several sterically or electronically comparable positions. But, later, it gained tremendous attention with the discovery of a prodigious number of natural products such as lignans,[2] xanthones[3] and a wide variety of tetrahydroisoquinoline alkaloids[4], which had been assumed[5–7] and later confirmed to be biosynthesized involving OPC of appropriate precursors. This reaction thus gained a new dimension by providing an excellent laboratory method for mimicking certain biosynthetic steps. This has initiated extensive studies on the reaction on a wide variety of phenolic compounds, leading to the successful biomimetic synthesis of many natural products, viz. gossypol analogue,[8] a well-known male antifertility agent (Scheme 1); euphorbetin and isoeuphorbetin[9], two coumarin dimers (Scheme 2); several alkaloids[10]; dimers



Gossypol analogue

Scheme 1.