Stabilizing interactions in the dimer interface of α -subunit in *Escherichia coli* RNA polymerase: A graph spectral and point mutation study

NATARAJAN KANNAN, PREETHI CHANDER, PALLAVI GHOSH,
SARASWATHI VISHVESHWARA, AND DIPANKAR CHATTERJI

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

(RECEIVED June 19, 2000; FINAL REVISION October 12, 2000; ACCEPTED October 19, 2000)

Abstract

The formation of α_2 dimer in *Escherichia coli* core RNA polymerase (RNAP) is thought to be the first step toward the assembly of the functional enzyme. A large number of evidences indicate that the α -subunit dimerizes through its N-terminal domain (NTD). The crystal structures of the α -subunit NTD and that of a homologous *Thermus aquaticus* core RNAP are known. To identify the stabilizing interactions in the dimer interface of the α -NTD of *E. coli* RNAP, we identified side-chain clusters by using the crystal structure coordinates of *E. coli* α -NTD. A graph spectral algorithm was used to identify side-chain clusters. This algorithm considers the global nonbonded side-chain interactions of the residues for the clustering procedure and is unique in identifying residues that make the largest number of interactions among the residues that form clusters in a very quantitative way. By using this algorithm, a nine-residue cluster consisting of polar and hydrophobic residues was identified in the subunit interface adjacent to the hydrophobic core. The residues forming the cluster are relatively rigid regions of the interface, as measured by the thermal factors of the residues. Most of the cluster residues in the *E. coli* enzyme were topologically and sequentially conserved in the *T. aquaticus* RNAP crystal structure. Residues 35F and 46I were predicted to be important in the stability of the α -dimer interface, with 35F forming the center of the cluster. The predictions were tested by isolating single-point mutants α -F35A and α -I46S on the dimer interface, which were found to disrupt dimerization. Thus, the identified cluster at the edge of the dimer interface seems to be a vital component in stabilizing the α -NTD.

Keywords: RNA polymerase; α -subunit; N-terminal domain; graph spectral algorithm; point mutation; assembly

Quantitative information on the strengths of the interchain interactions in oligomeric proteins is often imperative for a complete understanding of the quaternary structure of the protein and its function. The DNA-dependent RNA polymerase (RNAP) is one such multicomponent enzyme that catalyzes the transcription of DNA to RNA by using ribo-

nucleoside triphosphates. Core RNAP consists of four subunits, namely, 2α , β , and β' . The enzyme complex is assembled sequentially both in vivo and in vitro as $\alpha + \alpha \rightarrow \alpha_2 \rightarrow \alpha_2\beta \rightarrow \alpha_2\beta\beta'$ (Ishihama 1981). It appears that the dimerization of the α -subunit plays a pivotal role by triggering the enzyme assembly process, failing which β and β' subunits cannot assemble to form the core RNAP. Because the dimerization of the α -subunit is the first step in the assembly pathway (Kimura and Ishihama 1995), understanding the stabilizing interactions in the dimer interface of the two α -subunits is essential.

Reprint requests to: Dipankar Chatterji, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India; e-mail: dipankar@mbu.iisc.ernet.in; fax: 91-080-3600535.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.26201

It is now well established that the α -subunit of *E. coli* RNAP consists of two domains, where the N-terminal domain (NTD; 1–235 amino acids) is responsible for dimerization and the C-terminal domain (245–329 amino acids) is exclusively required for transcriptional activation (Igarashi and Ishihama 1991; Ebright and Busby 1995; Kimura and Ishihama 1995; Kimura and Ishihama 1996). A large number of mutations have been performed to identify the residues that are important in the dimerization procedure (Kimura and Ishihama 1995). However, except for the T38A mutation, no other point mutation has been found to be dimerization defective.

In the present study, we used a graph spectral algorithm (Kannan and Vishveshwara 1999) to identify residues that are important in the stability of the dimer interface by identifying side-chain clusters from the crystal structure of α -NTD of *E. coli* RNAP (Zhang and Darst 1998). Clusters of side-chain interactions are known to stabilize protein–protein complexes (Bogan and Thorn 1998). Identifying such clusters can help in identifying residues important for the structural stability of oligomeric proteins. A usual way of identifying such residues is to examine the spatial interactions of all the residues in the interface. However, following such a procedure may fail to identify residues located close to the interface that still affect the association between the subunits (Park et al. 1999).

In the graph spectral algorithm, the global nonbonded interactions of all the side chains in the protein molecule are represented in the form of a mathematical graph. This connected graph is represented in the form of a Laplacian matrix (see Materials and Methods). Clusters and cluster centers are obtained from the eigenvalues and eigenvectors of the Laplacian matrix. A nine-residue cluster adjacent to the hydrophobic core of the α -NTD homodimeric interface was identified by using this procedure. The cluster residues constitute both polar and hydrophobic interactions. The identified cluster is also topologically conserved in the *Taq* RNAP crystal structure. Two residues, F35 and I46, were identified as crucial in the stability of the α -dimer. Because the hydrophobic core in the dimer interface was thought to be the stabilizing unit (Zhang and Darst 1998), the free energy contribution (as measured by the change in accessible surface of the residues on dimer formation [Chothia and Janin 1975]) of the cluster and of the core residues was compared. Two point mutations that would be dimerization defective, namely, F35A and I46S, were predicted theoretically on the basis of the following three aspects: (1) the residues were identified to be important by the graph spectral method, (2) the residues were conserved among the *E. coli* and *Taq* sequences, and (3) the residues contributed highly to the change in accessible surface area (ASA) on dimer formation. Both of the mutants F35A and I46S, predicted on the basis of vector component values as well as from the conservation of the residues, showed a complete

disruption of the dimer formation. The present study highlights the importance of a specific cluster in stabilizing the dimer interface of the α -NTD domain and the efficacy of the graph spectral algorithm in identifying stabilization units.

Results and discussion

The α -NTD structure of *E. coli* RNAP is known at a resolution of 2.5 Å (Zhang and Darst 1998). The NTD-dimer constitutes two domains, wherein the first domain consists of four stranded antiparallel β sheets and two nearly orthogonal helices (H1 and H3), and the other domain consists of one helix and seven β strands with antiparallel arrangement (Fig. 1A). The dimer interface of the α -NTD is formed by the interaction between the two cognate helices H1 and H3 of the first domain from each monomer.

A graph spectral algorithm was used (see Materials and Methods) to identify side-chain clusters in the structure of α -NTD. A nine-residue cluster was identified in the dimer interface in which six residues, 8F, 35F, 31L, 32E, 38T, and 39L, emanate from one monomer (A chain) and three residues, 46I, 50S, and 227Q, emanate from the other monomer (B chain) (Fig. 1b).

Because the interface of the α -dimer is symmetric, an equivalent cluster is formed on the symmetrically opposite side of the dimer interface (Fig. 1B). Five of the nine cluster residues (35F, 38T, 39L, 46I, 227Q) emanate from the N-terminal region of helix H1 (A chain) and the C-terminal region of helix H1 and H3 (B chain). Three residues (E32, 8F, and 50S) are located in the loop regions connecting the helices to the strands, and L31 is located in the C-terminal region of strand 2 (Table 1). Side-chain clusters identified earlier on a large data set of proteins also showed that most of the side-chain clusters occur in the terminal regions of the helices and strands (Heringa and Argos 1991). The identified cluster in the α -NTD-dimer interface was further investigated for its structural location, stability, and residue composition.

Residue composition and cluster location

The identified nine-residue cluster in the dimer interface of the α -NTD constitutes both hydrophobic and polar residues (Table 1). A combination of hydrophobic and polar interactions are characteristic of subunit interfaces and are known to stabilize protein–protein interaction (Jones and Thornton 1997). A recent analysis by Larsen et al. (1998) of 136 homodimeric structures showed that a large number of homodimeric proteins are stabilized by a combination of hydrophobic and polar interactions. In the present case of the α -NTD homodimeric interface, the polar (E32, T38, S50, and Q227) and hydrophobic (35F, 8F, 31L, 39L, 46I) residues are found to form a cluster in the interface that

Fig. 1. (a) The location of the cluster in the dimer interface of α -N-terminal domain. (b) Enlarged view showing the cluster residues. The figure was created by using the package VMD (Humphrey et al. 1996).

would stabilize the interface more than would a mere pairwise polar–polar or hydrophobic–hydrophobic interaction.

Zhang and Darst (1998) identified an extended hydrophobic core in the interface of the α -NTD that is sandwiched between the two orthogonal pairs of antiparallel α helices (Fig. 2). In the present case, the identified cluster in the dimer interface is located adjacent to the hydrophobic core (Fig. 2). In fact, two residues of the hydrophobic core, I46 (B chain) and L39 (A chain), are also part of the identified

cluster. On reducing the side-chain interaction criteria, the nine-residue cluster expands by incorporating the hydrophobic core residues as part of an extended cluster.

The location of the cluster in the dimer interface was further investigated by calculating the percentage of ASA (Connolly 1993) of the cluster residues. The average percentage of ASA of the cluster was $5.50 \text{ \AA}^2$, with residues E32, T38, and S50 being partially exposed to the solvent (Table 1). This partial exposure of the cluster residues is

Table 1. Secondary structural location, thermal factors and accessible surface area for the identified cluster residues in the interface

Cluster residues (chain)	Secondary structural location ^a	Thermal factor ^b ($\text{\AA}^2$)	ASA ^c ($\text{\AA}^2$)
35F (A)	H1 (N-terminal)	20.19	3.753
8F (A)	L	16.39	15.73
31L (A)	S2 (C-terminal)	16.73	0
32E (A)	L	23.04	10.69
38T (A)	H1	10.17	9.77
39L (A)	H1	8.96	0
46I (B)	H1 (C-terminal)	12.41	0.82
50S (B)	L	18.43	0.37
227Q (B)	H3 (C-terminal)	17.60	0.28

^a (H) Helix; (S) strand; (L) loop; (ASA) percentage of accessible surface area.

^b Average thermal factor of the cluster residues is $23.24 \text{ \AA}^2$.

^c Average ASA of the cluster residues is $5.50 \text{ \AA}^2$.

consequent to its location at the edge of the dimer interface (Fig. 2).

Conservation of cluster residues

Recently, the crystal structure of the *Taq* core RNAP structure was solved to $3.3 \text{ \AA}$ by Zhang et al. (1999). The amino acid sequence of the α -NTD of *Taq* RNAP shows a percentage identity of 46.9% with the sequence from *E. coli*. Because the structures of the α -NTD in both *E. coli* and *Taq* are similar (Zhang et al. 1999), we identified clusters in the α -NTD of the *Taq* coordinates also. The identified cluster is found to be located in topologically similar locations in both structures, and the residues constituting the cluster are conserved in the amino acid sequences of *Taq* and *E. coli*, except residues F8, Q227, and S50 (Fig. 3). Both structural and sequential conservation of the cluster residues indicates that the identified cluster is important in the stabilization of the dimer interface.

Thermal factor

Stability of the identified cluster was assessed by using the crystallographic temperature factors that measure the flexibility of the thermal vibration state of individual protein atoms (Stout and Jensen 1968). The lower the temperature factor, the more rigid is the atom associated with it. The average temperature factor of all the side-chain atoms in the α -NTD structure and the temperature factor of the side-chains forming the nine-residue cluster in the dimer interface were calculated (Table 1). The thermal factor of all the side-chains forming the cluster is less than the average thermal factor of the protein, indicating structural rigidity of the cluster residues. This implicates the importance of cluster residues in stabilizing the dimer interface.

Prediction of hot spots on the dimer interface

Vector component magnitude

Cluster centers can be identified by examining the vector components corresponding to the highest eigenvalue. We had earlier shown that the residue with the largest vector component magnitude of the highest eigenvalue forms the center of the cluster (Patra and Vishveshwara 1998), and we pointed out a correlation between vector component magnitude of the cluster residues and the Φ_F values (a measure of the effect of mutation on the folding free energy) (Kannan and Vishveshwara 1999). In the present study, an examination of the vector component magnitude of the cluster residues shows the largest magnitude of 0.936 for 35F (Table 2), indicating this residue to be the center of the cluster. The residues 50S and 46I have the second and third largest vector component magnitude of 0.179 and 0.160 respectively. This indicates a strong contribution of these three residues to dimer stabilization.

Side-chain contribution to the dimer association

A measure of the loss in the ASA of two subunits when they associate together to form a dimer gives an estimate of the hydrophobic free energy of subunit association (Chothia and Janin 1975). A good estimate of this can be made when the structure of both the monomer and the dimer are known. However, in the absence of the monomer structure, it is reasonable to approximate the structure of the two monomers to be similar to that of the dimer. In the present study, the ASA of the α -NTD in the homodimeric form and that of the individual monomers was calculated by using Connolly's (1993) ASA program. The total ASA of the subunits separately was $21,435 \text{ \AA}^2$ and that of the complex was $18,300 \text{ \AA}^2$. The change in ASA on dimer formation is about $3000 \text{ \AA}^2$. An empirical correlation between the extent of ASA buried on dimerization and the hydrophobic free energy contribution has been given (Chothia and Janin 1975), according to which a burial of $1 \text{ \AA}^2$ corresponds to a contribution of 0.025 kcal/mole of hydrophobic free energy. Hence, a burial of $3000 \text{ \AA}^2$ of interface surface area would account for a hydrophobic free energy contribution of 75 kcal/mole .

Zhang and Darst (1998) identified some of the hydrophobic residues in the interface of α dimer as the "hydrophobic core" on the basis of the conservation of the hydrophobic residues among the α homologs in eukaryotes, bacteria, and chloroplasts. The hydrophobic core at the dimer interface is formed mostly by the residues emanating from the orthogonal pair of antiparallel helices H1 and H3 of each monomer.

The contribution of individual side chains of the cluster as identified by our method and that of the hydrophobic core to the change in ASA on subunit association is given in Table 3. The cluster residues contribute more to the burial of subunit interface compared with the hydrophobic core and

Table 2. Vector component magnitude of the cluster residues

Cluster residues	Vector component magnitude
35F (A chain)	0.936
50S (B chain)	0.179
46I (B chain)	0.160
8F (A chain)	0.155
227Q (B chain)	0.133
38T (A chain)	0.087
31L (A chain)	0.081
39L (A chain)	0.076
32E (A chain)	0.067

the residues in the cluster contributing the largest to the free energy of hydrophobic association (Table 3); and (3) conservation of the residues among the *E. coli* and *Taq* sequences (Fig. 3). It is important to note that predicting mutations just on the basis of the contribution of residues to the free energy of dimer formation or the conservation of residues may not be sufficient to identify specific residues that are important in dimer stabilization, because more than one residue shows a high Δ ASA on dimer formation (Table 3). Also, most of the hydrophobic residues in the interface are conserved (Fig. 3). Hence, using either of these methods or both methods together would not be sufficient to identify specific residues important in the dimer stabilization. However, using the graph spectral algorithm along with the other two procedures helps in identifying specific residues that are important in structure stabilization and also improves the confidence level of predicting mutations that would affect subunit association.

With these three procedures, it was found that residue 35F (chainA) satisfied all three conditions, followed by 46I from the other monomer (chain B). Residues that form the hot spots on the interface are known to be complementary and to pack against each other on the interface (Bogan and Thorn 1998). Hence, we predicted that two mutations, 35F (A chain) and 46I (B chain), would disrupt dimerization.

Mutation of predicted residues

Table 4 shows mutation sites that were created earlier to follow their role during the α dimerization step of RNAP assembly. In addition to the point mutations, there were deletions in NTD of the α -subunits that gave rise to an assembly-defective phenotype. It is apparent from Table 4 that most of the assembly-defective mutations arise from large deletions in α -subunits. On the other hand, point mutations in specific regions can be tolerated by the protein. In the cases where point mutations do not yield functional reconstitution of RNAP, it is mostly due to the defect in the interactions with the β or β' subunits during subsequent steps of RNAP assembly. Thus, it appears that the dimerization-defective mutants in α -subunit are rare. However, a

Table 3. Comparison of the change in accessible surface area of cluster and core residues on α dimer formation

Resides	Δ ASA ($\text{\AA}^2$)	
35F (A chain)	154.83	Cluster residues
8F (A chain)	175.74	
31L (A chain)	1.03	
32E (A chain)	7.41	
38T (A chain)	135.43	
39L (A chain)	88.18	
46I (B chain)	38.37	Core residues
50S (B chain)	52.66	
227Q (B chain)	228.78	
28L (A chain)	23.65	
39L (A chain)	88.21	
43L (A chain)	11.31	
46I (A chain)	38.37	
47L (A chain)	0.0	
51M (A chain)	0.0	
216A (A chain)	0.0	
217I (A chain)	17.69	
220A (A chain)	0.0	
221A (A chain)	81.16	
224L (A chain)	87.00	
228L (A chain)	162.29	Additional residues
231F (A chain)	252.17	
4S (A chain)	40.13	
2Q (A chain)	42.84	
52P (B chain)	18.01	
150R (B chain)	73.16	
219R (B chain)	52.10	
222T (B chain)	30.66	
223I (B chain)	17.19	
226E (B chain)	37.76	

(Δ ASA) Change in accessible surface area on dimer formation.

Total Δ ASA of cluster residues = 882.43 $\text{\AA}^2$.

Total Δ ASA of core residues from one monomer = 762.55 $\text{\AA}^2$.

point mutation of T38A was known to partially affect the dimer formation (Kimura and Ishihama 1995). Interestingly, T38 is part of the identified cluster and has a lesser vector component magnitude compared with 35F and 46I (Table 2). Hence, it is reasonable to assume that there will be no single amino acid that is not a part of the identified cluster at the α_2 -dimer interface that can be mutated to destroy protein-protein contact. So far, no attempts had been made to mutate residues 35F and 46I, and therefore we made the two point mutants F35A and I46S to study their effect on the assembly of α -subunits. It should be mentioned that there exist a few assembly-defective mutants of α -subunit, for example, α -D174A, which forms a partial dimer, but the further association of β or β' is not hindered (Kimura and Ishihama 1995). However, α -D174A is neither a part of the nine-residue cluster nor is it located in the dimer interface. In any case, we did not try to find out whether core mutants studied here can be induced to accept β or β' subunits. Impairment of the α_2 -dimer assembly on point mutation was our main focus. It should be mentioned at this point that

Table 4. Mutations created earlier to follow α -dimerization during RNA polymerase assembly

Mutation	α_2 Level ^a	References
α -R45A	+++	M. Kimura and A. Ishihama (1995)
α -L48A	+++	
α -L47A	++	-do-
α -N41A	++	
α -E80A	++	
α -L82A	++	
α -K86A	++	
α -E215A	++	
α -R44A	+	-do-
α -E76A	+	
α -I81A	+	
α -V173A	+	
α -L85a	+	-do-
α -T38A	+	
α -L201A	+	
α -P213A	+	
α -D174A	-	-do-
α -P179A	-	
α -D199A	-	
α - Δ N 20	++	M. Kimura, N. Fujita and A. Ishihama (1994)
α - Δ N 10	+	-do-
α - Δ N 30	+	
α - Δ N 40	+/-	-do-
α - Δ N 60	+/-	
α - Δ N 100	+/-	
α - Δ N 120	+/-	
α - Δ N 160	+/-	
α - Δ N 81-100	-	-do-
α - Δ N 121-140	-	
α - Δ N 181-200	-	
α - Δ N 201-220	-	
α - Δ N 221-240	-	
α - Δ N 161-180	+/-	-do-
α - Δ N 21-40	+/-	
α - Δ N 141-160	+/-	
α - Δ N 41-60	+/-	
α - Δ N 61-80	+/-	
α - Δ N 101-120	+/-	-do-
α (1-200)	+/-	
α (1-220)	+/-	-do-
α (21-235)	++	
α -H128R	++	
α -H132R	++	
α -H117RH128R	++	
α -H117RH132R	++	
α -H117RH128RH132R	++	
α - Δ 40	-	-do-
α - Δ (111-156)	-	-do-
α - Δ 230	++	-do-

Table 4. Continued

Mutation	α_2 Level ^a	References
α -C131A	++	S. Sujatha, A. Ishihama, and D. Chatterji (2000)
α -C54A	++	
α -C176A	++	
α -C54AC131A	++	

^a The level of α dimer was measured by gel-filtration HPLC. (+++) More than wild type; (++) wild-type level; (+) less than wild-type; (+/-) an intermediate peak in gel filtration; (-) undetectable.

during the growth of *E. coli* transformed with an overexpression plasmid-bearing mutant α -subunit, the growth of the organism is maintained with the help of normal chromosomal α , although its amount is minimal. However, on induction, production of mutant protein increases and any subsequent purification will have very little contamination from its chromosomal counterpart. Figure 4 shows the FPLC profile of wild-type α and its two mutants at the core of the dimerization interface. Interestingly, we could not detect any contamination from partial dimer or α_2 . This experimental observation gives support to the graph spectral approach of predicting important residues involved in the dimer formation and perhaps can be used as a rational method to select important residues involved in protein-protein interaction for further experimental investigations.

The graph spectral analysis here was performed with the crystal structure coordinates obtained from the α_2 dimer of *E. coli* and from the *Taq* RNAP. The consistency in results from either of the data sets indicates that β or β' assembly over α -dimer may not influence the core structure at the dimer interface (Kimura et al. 1994; Sujatha and Chatterji 1999; Sujatha et al. 2000).

Materials and methods

Cluster identification procedure

Side-chain clusters in the α -NTD of *E. coli* RNAP (PDB code 1bdf) were identified by using a graph spectral method, the detailed algorithm of which is described in a recent paper (Kannan and Vishveshwara 1999). By using this algorithm, we recently identified a variety of side-chain clusters in protein structures.

In brief, the graph spectral algorithm identifies side-chain clusters in protein structures by representing the nonbonded interaction of the side chains of the protein molecule in the form of a connected graph. The C_β atoms of the residues are considered as the vertices of the graph, and two vertices are connected by an edge if the side-chain interaction criterion of the two residues under consideration is satisfied. The connected graph is represented mathematically in the form of a Laplacian matrix. On diagonalizing the Laplacian matrix, clusters are identified from the vector components corresponding to the second-lowest eigenvalue, and cluster centers are identified from the vector components corresponding to the higher eigenvalues. We had earlier used a number of side-chain interaction criteria to identify a variety of clusters in protein struc-

Fig. 4. Gel-filtration FPLC of α point mutants. Purified derivatives were analyzed by Superdex 75 HR 10/30 (Pharmacia) gel filtration chromatography in 0.3M NaCl buffer. The elution profiles were recorded with an ultraviolet monitor at 280 nm. Elution position of marker proteins are indicated at the top: ovalbumin (43 kD), myoglobin (17.6 kD), and RNase A (13.7 kD).

tures. In the present study, a percentage side-chain interaction criterion of 5%, along with a high-contact criterion was used to identify clusters on the α -NTD structure. Side-chain interaction between residues R_i and R_j is evaluated on the basis of the number of side-chain atom pairs that come within an interatomic distance of 4.5 Å. Side-chain R_i is considered to satisfy a high-contact criterion if residue R_i interacts with more than three other side chains that are not sequentially adjacent. Use of high-contact criteria helps in the identification of a highly compact cluster of side chains.

Uniqueness of a graph spectral algorithm

A usual way of identifying clusters in a protein structure is to evaluate the "local" nonbonded interactions of the residues. Any set of residues that have a larger number of interactions among themselves than with the rest of the protein molecule is defined as a cluster (Heringa and Argos 1991). This way of identifying clusters considers only the local interactions of the residues (residues spatially close) for the clustering procedure, and the relative im-

portance of specific residues in the cluster is not available from this procedure. A few other methods available to identify clusters and their limitations are mentioned in our earlier article (Kannan and Vishveshwara 1999).

The described graph spectral algorithm in the present study considers the "global" nonbonded side-chain interactions of all the residues while identifying the residues that form a cluster. Moreover, the contribution of every side chain to the formation of the cluster and the residues that form the center of the cluster are identified by just examining the vector component magnitudes corresponding to the higher eigenvalues of the Laplacian matrix. No other method reported so far gives such quantitative, residue-wise information on the cluster residues, and the reported algorithm can be most appropriate in predicting stabilizing interactions by just examining the vector components corresponding to the higher eigenvalues of the cluster residues.

Mutagenesis of the *rpoA* gene

The *E. coli* RNAP α -subunit gene (*rpoA*) was cloned in an expression vector pGEMAX185 and was a kind gift from Professor Akira Ishihama. This gene was used as a template to generate site-directed mutants through PCR mutagenesis. Mutagenesis was performed by using the *QuickChange* protocol (Stratagene). The mutagenic oligonucleotides for the F35A mutation were GCCTT TAGAGCGTGGCGCTGGCCATACTCTG and CAGAGTATG GCCAGCGCCACGCTCTAAAGGC and those for the I46S mutation were CGCACTGCGCCGTAGTCTGCTCTCATCG and CGATGAGAGCAGACTACGGCGCAGTGCG. All mutant constructions were checked by DNA sequencing.

Expression and purification of the α -subunit

E. coli strain BL21 (DE3) was used for the overexpression of the wild-type α and the point mutants. Transformants were grown in LB medium containing 100–200 μ g/ml of ampicillin. Expression was induced at $OD_{600} = 0.6$ by adding 1 mM IPTG, and incubation was continued for 3 hr at 37°C; cells were harvested and stored at -80°C .

Proteins were purified according to Igarashi and Ishihama (1991). In brief, the proteins were extracted from the insoluble fraction with 0.1M NaCl in lysis buffer (50 mM Tris-HCl, pH 8.0, containing 1 mM EDTA), precipitated with 60% saturation of ammonium sulphate and resolubilized in 10 mM Tris-HCl (pH 7.6 at 4°C), 10 mM MgCl_2 , 0.1 mM EDTA, 0.1 mM dithiothreitol (DTT), 5% glycerol, and 0.1 M NaCl. After dialysis against the same buffer, the crude α -subunit was subjected to DE52 column chromatography. Proteins were eluted with a linear gradient of 0.1–0.5 M NaCl in the same buffer. Fractions containing the α -subunit were pooled and dialyzed against this buffer, containing 50% glycerol (storage buffer).

Gel filtration FPLC

The purified α -subunits were subjected to Superdex 75 HR 10/30 (Pharmacia) gel-filtration column chromatography at a flow rate of 0.4 ml/min, in a buffer containing 10 mM Tris-HCl (pH 7.6), 10 mM MgCl_2 , 0.1 mM EDTA, 0.1 mM DTT, 5% glycerol, and 0.3M NaCl, by using AKTA Basic FPLC system (Pharmacia). The elution profiles were recorded by monitoring the absorbance at 280

nm. Myoglobin, ovalbumin, and RNase were used as molecular mass markers.

Acknowledgments

The authors thank Akira Ishihama of National Institute of Genetics (Japan) for the overexpression *rpoA* gene and Seth Darst of Rockefeller University (New York) for the crystal structure coordinates of *Taq* RNAP. We also thank the Department of Biotechnology, Government of India for financial support to Dipankar Chatterji. The Super Computer Education and Research Center and the Bioinformatics Center at the Indian Institute of Science are acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bogan, A.A. and Thorn, K. 1998. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**: 1–9.
- Chothia, C. and Janin, J. 1975. Principles of protein-protein recognition. *Nature* **256**: 705–708.
- Connolly, M. 1993. The molecular surface package. *J. Mol. Graph.* **11**: 139–141.
- Ebright, R.H. and Busby, S. 1995. The *E. coli* RNA polymerase α subunit: Structure and function. *Curr. Opin. Genet. Dev.* **5**: 197–203.
- Heringa, J. and Argos, P. 1991. Side chain clusters in protein structures and their role in protein folding. *J. Mol. Biol.* **220**: 151–171.
- Humphrey, W., Dalke, A., and Schulten, K. 1996. VMD—Visual molecular dynamics. *J. Mol. Graph.* **14.1**: 33–38.
- Igarashi, K. and Ishihama, A. 1991. Bipartite functional map of the *E. coli* RNA polymerase α subunit: Involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **32**: 319–325.
- Ishihama, A. 1981. Subunit assembly of *Escherichia coli* RNA polymerase. *Adv. Biophys.* **14**: 1–35.
- Jones, S. and Thornton, J.M. 1997. Analysis of protein-protein interaction sites using surface patches. *J. Mol. Biol.* **272**: 121–132.
- Kannan, N. and Vishveshwara, S. 1999. Identification of side-chain clusters in protein structures by a graph spectral method. *J. Mol. Biol.* **292**: 441–464.
- Kimura, M. and Ishihama, A. 1995. Functional map of the α subunit of *E. coli* RNA polymerase: Amino acid substitution within the amino terminal assembly domain. *J. Mol. Biol.* **254**: 342–349.
- . 1996. Subunit assembly in vivo of *E. coli* RNA polymerase: Role of the amino terminal domain assembly domain of α subunit. *Genes Cells* **1**: 517–528.
- Kimura, J., Fujita, N., and Ishihama, A. 1994. Functional map of the α subunit of *E. coli* RNA polymerase. *J. Mol. Biol.* **242**: 107–115.
- Larsen, T.A., Olson, A.J., and Goodsell, D.S. 1998. Morphology of protein-protein interfaces. *Structure* **6**: 421–427.
- Park, Y.C., Guez, V., and Bedouelle, H. 1999. Experimental evolution of a dense cluster of residues in tyrosyl-tRNA synthetase: Quantitative effects on activity, stability and dimerisation. *J. Mol. Biol.* **286**: 563–577.
- Patra, S.M. and Vishveshwara, S. 1998. Classification of polymer structures by graph theory. *Int. J. Quantum Chem.* **71**: 349–356.
- Stout, G.H. and Jensen, L.H. 1968. *X-ray structure determination: A practical guide*. Macmillan, London.
- Sujatha, S. and Chatterji, D. 1999. Detection of putative Zn(II) binding sites within *E. coli* RNA polymerase: Inconsistency between sequence-based prediction and ^{65}Zn blotting. *FEBS Lett.* **454**: 169–171.
- Sujatha, S., Ishihama, A., and Chatterji, D. 2000. Functional complementation between mutations at two distant positions in *E. coli* RNA polymerase as revealed by second-site suppression. *Mol. Gen. Genet.* (in press).
- Zhang, G. and Darst, S.A. 1998. Structure of the *E. coli* RNA polymerase α subunit amino-terminal domain. *Science* **281**: 262–266.
- Zhang, G.E.A., Campbell, L.M., Richter, C., Severinov, K., and Darst, S.A. 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3Å resolution. *Cell* **98**: 811–824.