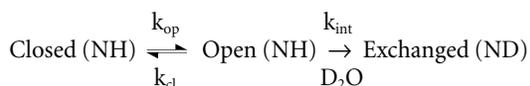


Fig. 2 Protection factors (P) and opening rates of the dimeric Lac HP62-V52C plotted as a function of residue number. Protection factors of the backbone amide protons of Lac headpiece in **a**, the free state (red bars) and in complex with the SymL (blue bar), **b**, wild type and **c**, SymR operators. In (**b**), orange bars refer to the residues of the left subunit, whereas the blue ones refer to the right. Protection factors were calculated from the ratio $k_{\text{int}} / k_{\text{obs}}$ and are displayed as a logarithmic scale. Values for residues 30 (right subunit) and 37 (left subunit) were not measured due to overlap. **d**, Opening rates, k_{op} , plotted as a function of residue number, of the backbone amide protons of the dimeric Lac HP62-V52C in complex with the SymL operator. All residues within a group cluster around the mean value, with maximum deviation of 15%. The color code is in accordance with that of Fig. 3c. The lowest measurable protection factor of free Lac headpiece is ~ 10 . Helix II and IV refer to the recognition and hinge helices, respectively.

bonds or burial in the binding interface. The exchange rates were analyzed according to the following scheme¹⁸



According to the model, exchange can take place only from the open conformation, with a rate constant k_{int} , which depends on sequence, pH and temperature and can be easily calculated¹⁹. Open and closed conformations interconvert with rate constants k_{op} and k_{cl} , respectively. There are two mechanisms by which exchange can take place: EX2 and EX1. In the first case, the rate

constant for re-protection, k_{cl} is much greater than k_{int} and the observed hydrogen exchange reflects the equilibrium constant between the closed and the open states: $k_{\text{obs}} = (k_{\text{op}} / k_{\text{cl}}) k_{\text{int}}$. The ratio $k_{\text{cl}} / k_{\text{op}}$ is referred to as the amide protection factor and can be used to estimate the free energy for the dominant opening reaction $\Delta G_{\text{op}} = -RT \ln (k_{\text{obs}} / k_{\text{int}})$. At the other extreme, when k_{int} is much greater than the rate constant for re-protection (typically above pH 8–9), the mechanism becomes EX1; the observed rate constant becomes independent of k_{int} and simply equals the rate constant for the formation of the unprotected state, k_{op} . Thus, exchange rates measured under both conditions can be used to extract both the thermodynamics and kinetics of the opening event. Recently, this simple model for NH exchange had been shown to be a robust framework for obtaining quantitative information about molecular motions in native proteins²⁰.

DNA-induced redistribution of the native-state ensemble

Protection factors for Lac HP62-V52C in the free state (Fig. 2a, red bars) are unexpectedly low (average value $< 10^2$) for a folded protein, which typically has values ranging from 10^4 to 10^8 . We were able to measure exchange rates only at low temperature (290 K) and pH (pD = 4.5). The low stability of the molecule in the uncomplexed state is further corroborated by thermal denaturation experiments, which showed significant unfolding of the α -helices already at 42 °C (data not shown). However, upon binding to the SymL operator (Fig. 1b), there is a dramatic increase in protection, with maximum values up to

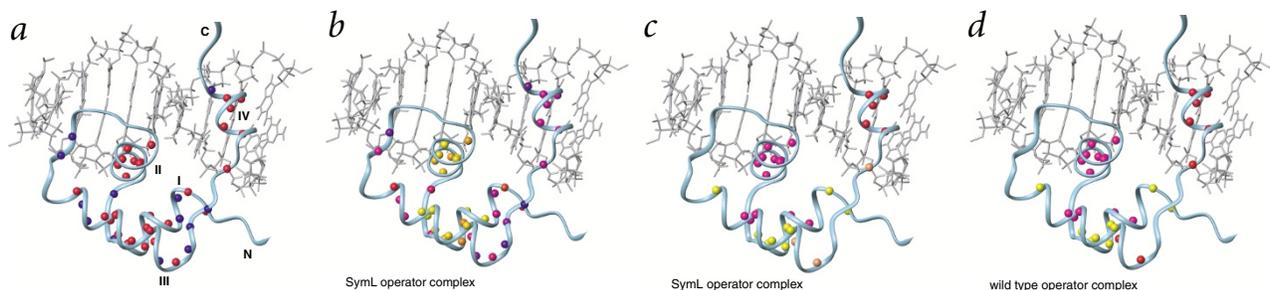


Fig. 3 Summary of the thermodynamic and kinetic data extracted from hydrogen exchange data. For clarity reasons, only the left site of the complex is depicted. Helix II and IV refer to the recognition and hinge helices, respectively. **a**, All backbone amide protons that are protected in the DNA-bound state. NHs that exchange only under the EX2 mechanism are blue, whereas those that switch to EX1 at high pH values are red. **b**, Difference in the free energy of the opening process, $\Delta\Delta G$, at each residue site between the free and SymL-bound state. Values of $\Delta\Delta G$ (kcal mol⁻¹) are displayed with a continuous color scale (6 (blue)–12 kcal mol⁻¹ (yellow)). **c**, Mapping of the opening rates, k_{op} , of the dimeric Lac HP62-V52C bound to SymL operator. The color code used to display k_{op} is as follows: $k_{\text{op}} = 0.20 \text{ h}^{-1}$ is red; $k_{\text{op}} = 0.11 \text{ h}^{-1}$, orange; $k_{\text{op}} = 0.04 \text{ h}^{-1}$, yellow; and $k_{\text{op}} = 0.02 \text{ h}^{-1}$, violet (Fig. 2d). **d**, Mapping of the opening rates, k_{op} , of the dimeric Lac HP62-V52C bound to wild type operator. Only the left site is depicted because the residues of the right subunit show the same rates with the exception of Asn 50, which is not protected. The color code used to display k_{op} is as follows: $k_{\text{op}} = 1.9 \text{ h}^{-1}$ is red; $k_{\text{op}} = 0.75 \text{ h}^{-1}$, yellow; and $k_{\text{op}} = 0.39 \text{ h}^{-1}$, violet. Note that the same color in panels (c,d) do not mean the same rates. The data are projected on the three-dimensional structure of the Lac HP62 complex with the SymL operator⁷. All models were drawn in MOLMOL²⁹ using Protein Data Bank entry 1CJG.



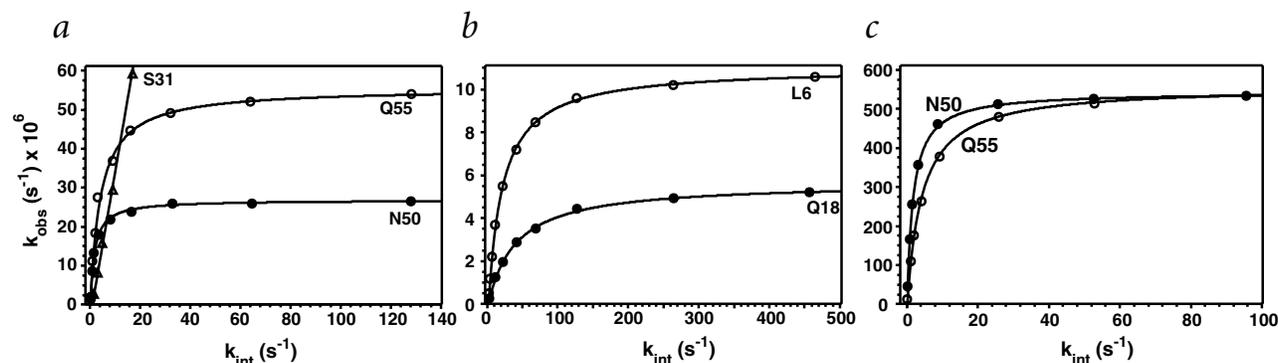


Fig. 4 Plot of equilibrium (observed) exchange rate, k_{obs} , versus intrinsic exchange rate, k_{int} . The behavior of selected residues of the dimeric Lac HP62-V52C complex with **a**, **b** SymL and **c**, wild type operator is illustrated. Exchange under EX2 conditions shows a linear dependence on pH and the k_{int} rate increase by a factor of ten for each pH unit. The switch to EX1 mechanism occurs at the pH where $k_{\text{cl}} = k_{\text{int}}$, in which case the plateau value of k_{obs} approaches the opening rate. Therefore, residues with slower reprotection rates are expected to switch to EX1 at lower pH. Curves are fit to the two-state model for exchange, according to the equation $k_{\text{obs}} = k_{\text{op}} \times k_{\text{int}} / (k_{\text{cl}} + k_{\text{int}})$. As reported, the stability of the Lac repressor–operator complex is rather insensitive to pH³⁰.

10^8 (Fig. 2a, blue bars). Even more impressive is the observation that protection is not confined to the binding site but is spread throughout the whole molecule (Figs 2a, 3b), as indicated through locally resolved free energy changes in protein upon DNA binding (Fig. 3b). This clearly means that there is a profound reduction in the various conformational states from which exchange can occur.

The natural *lac* operator (Fig. 1a) is asymmetric, and the two binding sites differ significantly in their affinity for Lac repressor when considered separately²¹. However, measured protection factors for the dimeric Lac headpiece in complex with its cognate operator are the same for the left and the right protein subunit (Fig. 1b). The only difference is Asn 50, which hydrogen bonds to DNA only in the left site. In order to probe the cooperative effect of hinge helices destabilization on the redistribution of the native-state ensemble, we used the complex with the SymR operator (Fig. 1c), in which the hinge helices exist at equilibrium between α -helical and random coil conformation¹⁰. The measured hydrogen exchange rates show no protection in the hinge region (Fig. 2c), whereas the conformational states from which exchange takes place are clearly more populated compared to the complexes with SymL and wild type operator. The difference in the reduction in the manifold conformational states upon Lac binding to DNA may explain the variation of heat capacity reported earlier for various operators²².

Dissociation of the protein–DNA complex

To follow the dissociation of the protein–DNA complex with residue-level specificity, we measured hydrogen exchange of the dimeric Lac HP62-V52C bound to SymL operator as a function of pH. For local exchange, EX2 kinetics are generally observed as expected, because the native state fluctuations are rapid and there is a fast closing step (Fig. 4a). However, most of the residues switched to the EX1 mechanism at elevated pH values (Figs 3a, 4a,b). The possibility of following the kinetics of structural fluctuations required to break the hydrogen bond and/or expose individual amides to solvent is of great importance because it may reveal concerted phenomena that cannot be followed with conventional experiments. Opening rates segregate in four distinct groups (Fig. 2d). As shown mapped onto a three-dimensional structure of the complex (Fig. 3c), the first group to open is the hinge helix, with an exchange rate of 0.20 h^{-1} , fol-

lowed by Asn 50 and the C-terminal residues of the third helix (0.11 h^{-1}). Unfolding of this part of the protein is propagated to the rest of the molecule by progressive destabilization and collapse of protein–protein and protein–DNA contacts, resulting in the complex being dissociated with a rate of 0.02 h^{-1} . The apparent dissociation constant of the complex, measured by biochemical methods¹⁰, is $\sim 0.05 \text{ h}^{-1}$, indicating that the slowest k_{op} rate corresponds to the macroscopic kinetics of dissociation. The group of backbone amide hydrogens that are slowest to exchange tend to cluster in mutually packed elements of secondary structure, suggesting the existence of a partially folded submolecular domain. The structure of this domain is stabilized by the specific contacts to DNA from the recognition helix, which remains intact, and by strong hydrophobic interactions with residues from the other helices (Fig. 3c). Based on the variation of the lifetime of protected states, a model of progressive unfolding and dissociation of the protein–DNA complex is proposed (Fig. 5). Although identical k_{op} rates do not necessarily imply that cooperative unfolding takes place²³, binding of Lac repressor to DNA is accompanied by cooperative phenomena. For example, formation of the hinge helix, which has been shown to require both protein–protein and protein–DNA contacts, is stable only when the hinge helix of the other subunit is folded^{8,10}. The recognition helix should be the last one to dissociate from DNA, because the side chains of its residues form most of the contacts to the operator. Measurement of individual opening rates can reveal the pathway of the progressive unfolding of subglobal structural units, providing a clear picture of how the protein dissociates from DNA.

In view of the asymmetry of the two binding sites, we also investigated the dissociation pathway of Lac repressor from its wild type operator. Does the least stable site (the right one) open more frequently? Exchange rate measurement as a function of pH allowed us to follow the kinetics of the dissociation under EX1 conditions (Fig. 4c). The opening rates are the same for both protein subunits. As compared with the SymL complex, hinge helices open with a ~ 10 -fold higher rate, whereas the complex dissociates $\sim 20\times$ faster (0.39 h^{-1}) (Fig. 3d). Overall, the complex responds to hinge helix unfolding twice as fast as the SymL complex, but the pathway for dissociation remains the same (Fig. 3d). Although the two sites of the cognate operator show different affinity for Lac repressor when considered inde-

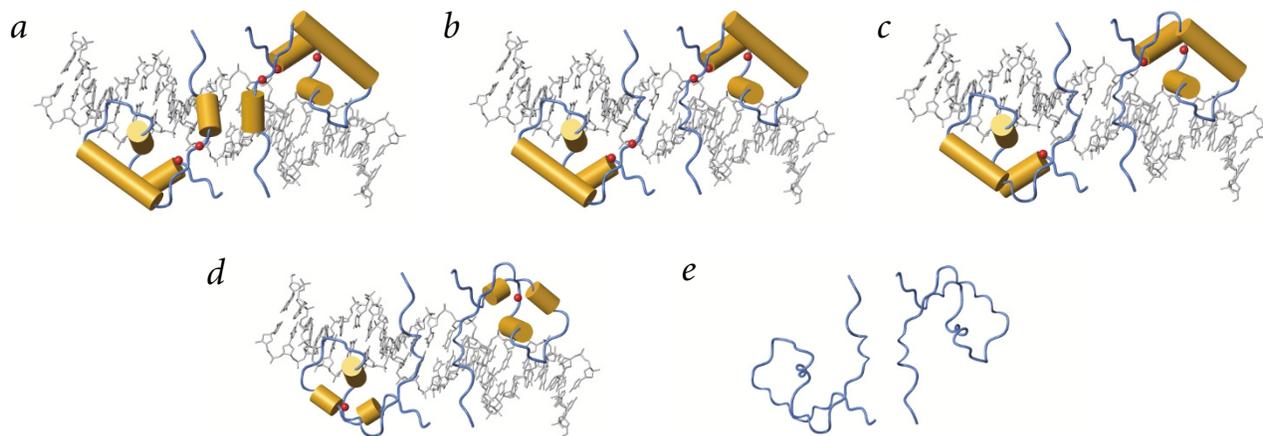


Fig. 5 Model of the dissociation pathway of the Lac repressor from the ideal SymL operator. The progressive unfolding of individual structural units is based on the opening rates measured by hydrogen-deuterium exchange for the individual sites. Along with the helices, three backbone amides are displayed (as spheres) that form crucial hydrogen bonds. Leu 6 and Asn 50 hydrogen bond to DNA phosphate, whereas the backbone amide of Val 15 hydrogen bonds to the carbonyl oxygen of Ala 10 and is a key residue in defining the proper orientation of the HTH domain. **a**, All rarely opened residues are protected. **b**, The first structural unit to unfold is that of the hinge helices, **c**, followed by disruption of the Asn 50 hydrogen bond to DNA and opening of the C-terminal residues of the third helix. **d**, Further unfolding takes place in which the hydrogen bond of Leu 6 to DNA disrupts and the only helix that remains intact is the recognition helix. A slowest to exchange core remains, which is stabilized by specific contacts to DNA from the recognition helix and strong hydrophobic interactions with residues from the other helices. Val 15 and the residues located in the middle of the first and third helices do not contact DNA, but are apparently important because they are involved in critical packing interaction with the recognition helix that stabilizes this submolecular domain. **e**, Their disruption results in protein–DNA dissociation.

pendently, their behavior is the same in the full operator by virtue of the cooperative nature of the hinge helices. Attempts to measure exchange rates of the dimeric Lac headpiece bound to the SymR operator under conditions that favor EX1 mechanism failed due to very fast exchange. Hinge helix destabilization forces the SymR complex to dissociate much faster than that of SymL, and there is an overall decrease in binding affinity by a factor of $\sim 1,000$ (Fig. 2c)²⁴.

Insight into the protein–DNA association pathway

The approach we followed in the present work could reveal important information about the association pathway of the protein–DNA complex as well. The key to accomplishing this is to measure the closing rates, specifically the rates by which amide protons become reprotected (see Methods). If residues that are protected only in the DNA-bound state are considered, the potential exists for monitoring the process of reprotection of individual structural units. The analyses show that Gln 18, a crucial residue of the recognition helix, becomes solvent protected approximately two times faster than Leu 6, with a rate of $\sim 40 \text{ s}^{-1}$ (this is consistent with an apparent association rate constant of $\sim 10^5 \text{ M}^{-1} \text{ s}^{-1}$). In conjunction with the observation of the submolecular core that is slowest to exchange, the results suggest that the recognition helix is the first to fit to DNA. This orientation is then locked by subsequent hydrogen bond formation of Leu 6 backbone to DNA. This contact is highly conserved among helix–turn–helix (HTH) proteins, irrespective of the nature of the residue²⁵. Hinge helix formation is slower than Gln 18 protection by a factor of ~ 10 , and the last residue to close is Asn 50. This observation suggests that protein–DNA interactions in the minor groove are difficult to establish and can take place only after the HTH core of the headpiece has properly oriented with respect to DNA. Thus, the DNA sequence is first read out in the major groove by the recognition helix, followed by discrimination of the minor groove by the hinge region.

Implications for the allosteric mechanism and DNA recognition

The significant instability of the DNA-binding domain in the uncomplexed state appears to have severe implications for the allosteric mechanism whereby intact Lac repressor exerts control over gene expression. If only ‘high stability’ residues constituted the binding site, then all the states in the ensemble would be binding competent, and DNA binding would induce only an energy shift without any redistribution of the states in the ensemble¹. In that case, unfolding of the hinge helices would affect only binding to the minor groove. The dramatic difference in mobility between the free and DNA-bound Lac repressor can be used advantageously by the biological system to switch between functional states. Therefore, alteration of dynamics between free and bound state coupled to destabilization of crucial structural subunits may provide a level of control that allows rapid and accurate response of the biological system to changes in the cell environment³. Furthermore, the present results implicate DNA as an allosteric activator of Lac repressor^{26,27}. Another functional advantage of the high flexibility of Lac repressor in the free state is the ability to recognize various operator sequences.

As demonstrated here, two important phenomena accompany recognition of sequence-specific operators by Lac repressor: hinge helix formation and a profound reduction in the protein native-state ensemble. The flexibility of the DNA-binding domain of the Lac repressor in the presence of noncognate operator sites allows it to slide along the DNA, facilitating target location. Once the specific site has been recognized in the major groove, the hinge helices form and recognize the minor groove. This recognition is followed by the selection of only the binding-competent states of the ensemble, which manifests itself as a negative contribution to the heat capacity of the system.

The present results provide a unique view of the order of events that result in protein–DNA association and dissociation, enriching our understanding of how complex systems function.



Moreover, they present a compelling example of a large redistribution in the ensemble of conformational states of a protein when it binds specifically to DNA.

Methods

Sample preparation. Uniformly ^{15}N -labeled dimeric Lac-HP62-V52C was prepared as described¹⁰. All Lac operator DNA fragments (Fig. 1) were purchased at Carl Roth GmbH and further purified on a Q-Sepharose (Pharmacia) column. Free dimeric Lac HP62-V52C was dissolved in buffer containing 60 mM KPi and 400 mM KCl, whereas the complexes with DNA were dissolved in 10 mM KPi and 20 mM KCl. Trace amounts of NaN_3 were added as a preservative. Protein concentration in all samples was 1 mM, and the protein:DNA ratio was adjusted to 1:1 by titration.

NMR spectroscopy. All spectra were collected on Bruker DRX spectrometers equipped with triple-resonance probeheads operating at 750 and 600 MHz (^1H frequency). A typical HSQC spectrum was recorded with four scans per free induction decay (FID), 64 complex ^{15}N points and 1,024 complex ^1H points. For the complex with wild type operator, the spectra were recorded with higher resolution in the ^{15}N dimension (128 or 256 complex points). Experiments for the dimeric Lac HP62-V52C in the free state were performed at 290 K and pD = 4.5. Exchange rates for the complex with SymL and wild type operators were determined for numerous pD values, in the range 5.1–9.7, at 320 K. Due to low stability of the complex with SymR operator, experiments were performed at 303 K and pD = 5.5 for this system.

Determination and analysis of exchange rates. Amide proton exchange rates were determined from the time course of the peak intensities in a series of ^1H - ^{15}N heteronuclear single quantum correlation (HSQC) spectra after dissolving lyophilized samples in D_2O . The final pD was adjusted by DCI or NaOD solutions and was measured with a glass electrode at the temperature of the exchange experiment, taking into account the isotope effect: $\text{pD} = \text{pH}_{\text{read}} + 0.4$. Exchange rates were determined by plotting the intensities of each residue against time and fit by a first-order rate expression. Rates are accurately determined, and the typical error for data fitting is 5%. Determination of closing rates, k_{cl} , is less accurate as compared to opening rates because fitted values of k_{cl} depend entirely on knowledge of k_{int} . In this respect, predicted values of k_{int}

generally agree to within a factor of three with rates observed in denatured proteins, and most of the measured and predicted values are within a factor of two²⁸. The systematic measurement of exchange under conditions where both equilibria and opening rates are assessed allows us to estimate relatively accurate k_{cl} rates.

Acknowledgments

This work was financially supported by the Netherlands Foundation for Scientific Research (NOW-CW)

Correspondence should be addressed to R.K. email: kaptein@nmr.chem.uu.nl

Received 26 September, 2001; accepted 15 January, 2002.

- Freire, E. *Proc. Natl. Acad. Sci. USA* **96**, 10118–10122 (1999).
- Dunker, A.K. *et al. J. Mol. Graph. Model.* **19**, 26–59 (2001).
- Wright, P.E. & Dyson, H.J. *J. Mol. Biol.* **293**, 321–331 (1999).
- Lefstin J.A. & Yamamoto, K.R. *Nature* **392**, 885–888 (1998).
- Bell, C.E. & Lewis, M. *Curr. Opin. Struct. Biol.* **11**, 19–25 (2001).
- Lewis, M. *et al. Science* **271**, 1247–1254 (1996).
- Spronk, C.A.E.M. *et al. Structure* **7**, 1483–1492 (1999).
- Spronk, C.A.E.M., Slijper, M., van Boom, J.H., Kaptein, R. & Boelens, R. *Nature Struct. Biol.* **3**, 916–919 (1996).
- Bell, C.E. & Lewis, M. *Nature Struct. Biol.* **7**, 209–214 (2000).
- Kalodimos, C.G., Folkers, G.E., Boelens, R. & Kaptein, R. *Proc. Natl. Acad. Sci. USA* **98**, 6039–6044 (2001).
- Englander, S.W. & Krishna, M.M.G. *Nature Struct. Biol.* **8**, 741–742 (2001).
- Englander, S.W. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238 (2000).
- Xu, Y., Mayne, L. & Englander, S.W. *Nature Struct. Biol.* **5**, 774–778 (1998).
- Chamberlain, A.K., Handel, T.M. & Marqusee, S. *Nature Struct. Biol.* **3**, 782–787 (1996).
- Kiefhaber, T. & Baldwin, R.L. *Proc. Natl. Acad. Sci. USA* **92**, 2657–2661 (1995).
- Zahn, R., Perrett, S., Stenberg, G. & Fersht, A.R. *Science* **271**, 642–645 (1996).
- Parker, M.J., Dempsey, C.E., Hosszu, L.L.P., Waltho, J.P. & Clarke, A.R. *Nature Struct. Biol.* **5**, 194–198 (1998).
- Hvidt, A. & Nielsen, S.O. *Adv. Protein Chem.* **21**, 287–386 (1966).
- Bai, Y., Milne, J.S., Mayne, L. & Englander, S.W. *Proteins* **17**, 75–86 (1993).
- Sivaraman, T., Arrington, C.B. & Robertson, A.D. *Nature Struct. Biol.* **8**, 331–333 (2001).
- Sasmor, H.M. & Betz, J.L. *Gene* **89**, 1–6 (1990).
- Spolar, R.S. & Record, M.T. Jr. *Science* **263**, 777–784 (1994).
- Arrington, C.B. & Robertson, A.D. *J. Mol. Biol.* **300**, 221–232 (2000).
- Falcon, C.M. & Matthews, K.S. *Biochemistry* **39**, 11074–11083 (2000).
- Wintjens, R. & Rooman, M. *J. Mol. Biol.* **262**, 294–313 (1996).
- Ciubotatu, M., Bright, F.V., Ingersoll, C.M. & Koudelka, G.B. *J. Mol. Biol.* **294**, 859–873 (1999).
- van Tilborg, M.A. *et al. J. Mol. Biol.* **301**, 947–958 (2000).
- Arrington, C.B. & Robertson, A.D. *Biochemistry* **36**, 8686–8691 (1997).
- Koradi, R., Billeter, M. & Wüthrich, K. *J. Mol. Graph.* **14**, 52–55 (1996).
- Riggs, A.D., Bourgeois, S. & Cohn, M. *J. Mol. Biol.* **53**, 401–417 (1970).