Identification of critical residues in the colicin E9 DNase binding region of the Im9 protein

Michael J. OSBORNE*, Russell WALLIS†, Kit-Yi LEUNG†, Glyn WILLIAMS‡, Lu-Yun LIAN§, Richard JAMES†, Colin KLEANTHOUS† and Geoffrey R. MOORE*||

Schools of *Chemical and †Biological Sciences, University of East Anglia, Norwich NR4 7TJ, U.K., ‡Roche Discovery Welwyn, 40 Broadwater Road, Welwyn Garden City, Herts, AL7 3AY, U.K., and §Biological NMR Centre, P.O. Box 138, Maurice Shock Building, Leicester University, Leicester LE1 8HN, U.K.

^1H – ^{15}N NMR studies, in conjunction with mutagenesis experiments, have been used to delineate the DNase-binding surface of the colicin E9 inhibitor protein Im9 (where Im stands for immunity protein). Complexes were formed between the 15 kDa unlabelled E9 DNase domain and the 9.5 kDa Im9 protein uniformly labelled with ^{15}N . Approx. 90% of the amide resonances of the bound Im9 were assigned and spectral parameters obtained from ^1H – ^{15}N heteronuclear single quantum coherence (HSQC) spectra were compared with those for the free Im9 assigned previously. Many of the amide resonances were shifted on complex formation, some by more than 2 p.p.m. in the ^{15}N dimension and more than 0.5 p.p.m. in the ^1H dimension. Most of the strongly shifted amides are located on the surfaces of two of the four helices, helix II and helix III. Whereas helix II had already been identified through genetic and biochemical investigations as an important determinant of biological specificity, helix III had not previously been implicated in binding

to the DNase. To test the robustness of the NMR-delineated DNase-binding site, a selection of Im9 alanine mutants were constructed and their dissociation rate constants from E9 DNase-immunity protein complexes quantified by radioactive subunit exchange kinetics. Their off-rates correlated well with the NMR perturbation analysis; for example, residues that were highly perturbed in HSQC experiments, such as residues 34 (helix II) and 54 (helix III), had a marked effect on the DNase-immunity protein dissociation rate when replaced by alanine. The NMR and mutagenesis data are consistent with a DNase-binding region on Im9 composed of invariant residues in helix III and variable residues in helix II. The relationship of this binding site model to the wide range of affinities (K_d values in the range 10^{-4} to 10^{-16} M) that have been measured for cognate and non-cognate colicin DNase-immunity protein interactions is discussed.

INTRODUCTION

There is an extensive literature concerning the structural and thermodynamic basis for specificity in protein–DNA interactions [1–3] but in contrast relatively little is known about the rules that govern selectivity in protein recognition. In recent years our knowledge in this area has expanded as complexes involving antibodies with protein antigens [4–6], ribonucleases with their inhibitors [7–10] and hormones with their receptors [11–13] have become almost as well characterized as proteinase-inhibitor complexes [14]. However, even with this information base only limited rules concerning the specificity of protein–protein interactions are discernible.

The highly specific colicin–immunity protein system is a novel and amenable model for addressing the principles that underlie selectivity in protein–protein recognition [15], particularly as the DNase colicins and their inhibitor proteins constitute a family of closely related natural mutant proteins in which specificity both *in vivo* and *in vitro* can be analysed [16]. These proteins also have relatively small molecular masses and can thus be studied by NMR. In the present study the interaction of the 15 kDa DNase domain of colicin E9 with its 9.5 kDa cognate inhibitor protein, Im9, has been studied by NMR to identify potential residues in the interprotein interaction site on the inhibitor. This study has been coupled with kinetic measurements of the DNase interacting with mutants of Im9 which were constructed by site-directed mutagenesis using the NMR data as a guide.

Colicin E9 (ColE9) is a member of the E colicin family of bacterial toxins, which utilize the BtuB receptor to enter bacterial cells and, following translocation across the cell membrane, initiate cell death [15]. Colicin-producing organisms avoid suicide by the co-ordinate synthesis of an acidic immunity protein that binds to the basic cytotoxic domain of the colicin and neutralizes the bacteriocidal activity within the producing host. There is extremely tight binding between cognate colicins and immunity proteins, as illustrated by the interaction of the DNase ColE9 with its inhibitor protein, Im9, which has a dissociation constant of approx. 10^{-16} M [17]. The colicin–immunity protein interaction is highly specific because unique inhibitor proteins Im2, Im7, Im8 and Im9 are required to provide the producing cells with complete protection against the action of their cognate colicins E2, E7, E8 and E9. Despite the high degree of sequence identity between the four DNase cytotoxic domains (approx. 80%) and their corresponding four immunity proteins (approx. 50%), cross-protection between non-cognate partners is readily observed only when the immunity proteins are overproduced in bacterial cells, with the cross-reactivities for Im2, Im7 and Im8 towards ColE9 mirroring the affinities *in vitro* for the E9 DNase (K_d 10^{-4} to 10^{-8} M) [16].

To define the origin of the specificity between ColE9 and its cognate inhibitor Im9, structural data are required for the proteins in their free and complexed forms. We have overproduced both the 134-residue DNase domain of ColE9 (the E9 DNase) [18] and the 86-residue Im9 [19], and used NMR to

Abbreviations used: Col, colicin; Im9, colicin E9 immunity protein; 2D, two-dimensional; 3D, three-dimensional; HMQC, heteronuclear multiple quantum coherence; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement.

|| To whom correspondence should be addressed.

determine the structure of Im9, which can be described as a distorted four- α -helical bundle [20,21]. Additionally, homologous recombination experiments have identified the region of Im9 between residues 16 and 43, including helix II, which encompasses residues 30–43 [20], as being important for specificity [16]. Independently, Chak et al. [22] have determined the structure of the homologous colicin E7 immunity protein by X-ray crystallography and have proposed that helices I and II constitute the DNase-binding region. In the present paper we present ^1H – ^{15}N NMR data that define the surface of Im9 involved in the interaction with the E9 DNase. The data show that in addition to helix II a region that is homologous throughout the DNase immunity protein family is also important for binding. Our binding model was tested by mutagenesis experiments, in which five residues within this surface were replaced with alanine and the dissociation rate constants for these complexes were compared with that of the wild-type Im9 protein. The results of this work are inconsistent with the proposals based on the X-ray structure determination of Im7 [22].

MATERIALS AND METHODS

Im9 and DNase

Samples of the E9 DNase domain and ^{15}N -labelled Im9 were prepared as described previously [18,20]. All samples of the E9 DNase–Im9 complex used for NMR contained approx. 1.2 mM Im9 and approx. 1.3 mM DNase in 20 mM phosphate buffer, $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ (9:1), pH 6.2.

NMR spectroscopy

Spectra were acquired with a Bruker AMX 600 spectrometer. ^1H and ^{15}N chemical shift referencing was performed as described by Osborne et al. [20]. All the following data were collected with a ^1H (F_3) carrier frequency at 4.80 p.p.m., a ^1H (F_3) spectral width of 8474 Hz, a ^{15}N (F_2) carrier frequency of 98.21 p.p.m. and a ^{15}N (F_3) spectral width of 2200 Hz. All multidimensional spectra were acquired in the pure phase absorption mode by using the time-proportional phase incrementation method described by Marion and Wüthrich [23]. For the two-dimensional (2D) and three-dimensional (3D) heteronuclear experiments, 32 and 165 transients respectively were collected; dataset sizes were $1024 (F_1) \times 128 (F_2)$ complex points for the 2D experiments and $1024 (F_1) \times 32 (F_2) \times 100 (F_3)$ complex points for the 3D experiments. For experiments requiring Hartmann–Hahn spin-lock, an MLEV 17 (2D homonuclear Hartmann–Hahn) [24,25] or a DIPSI-II [3D ^{15}N – ^1H heteronuclear multiple quantum coherence (HMQC)–TOCSY] [26,27] mixing pulse with total mixing times of 35 ms were used. For the 2D ^{15}N – ^1H correlation spectrum, either a heteronuclear single quantum coherence (HSQC) [28] or an HMQC [29] pulse sequence was used, with a GARP (globally optimized alternating rectangular pulses) sequence to decouple ^{15}N from protons during acquisition. We found that for the free Im9 protein both spectra gave similar sensitivities but for the complex spectrum the HMQC experiment was preferred. In the heteronuclear experiment, the water signal was suppressed by using a 2.5 ms gradient pulse of gradient strength 75 G/cm ($1 \text{ G} = 10^{-4} \text{ T}$).

Mutagenesis of *imm9*

Single-site alanine mutations were introduced into the *imm9* gene (encoding Im9) by the megaprimer method [30]. Mutated genes were cloned into the expression vectors pTrec99A or pKK233-2 (Pharmacia) and introduced into the host *Escherichia coli*

strain JM83 by transformation. In each case, the site-directed mutation and the entire immunity gene sequence was verified by double-stranded DNA sequencing with an A.L.F. sequencer (Pharmacia) and the reagents and protocols supplied by Pharmacia. The expression and purification of alanine mutants were as described previously for wild-type Im9 [19].

Determining values for the dissociation rate constant, k_{off}

Complex dissociation rates were measured by subunit exchange as described by Wallis et al. [17] except that ^3H Im9 was exchanged into unlabelled complexes. The amount of radioactivity in protein complexes was determined by scintillation counting after removal of the unbound ^3H Im9 by gel-filtration chromatography.

RESULTS AND DISCUSSION

NMR spectra of Im9 bound to the E9 DNase

The 600 MHz ^1H – ^{15}N HSQC spectra of uniformly ^{15}N -labelled Im9 free and bound to the unlabelled cytotoxic domain of ColE9 are shown in Figure 1. The DNase sample was not isotopically enriched with ^{15}N and does not contribute to the spectrum. Solution conditions for the two spectra were the same; therefore differences between them are a result of the formation of the DNase–Im9 complex. A slight excess of DNase over Im9 was used to ensure that the Im9 was fully bound. There is a very low rate of dissociation for the complex in the absence of salt, $t_{1/2} > 500 \text{ h}$ [17], and therefore exchange between free and bound Im9 is not a complication in the NMR analysis presented below.

Comparison of ^1H – ^{15}N HSQC spectra of free Im9 (Figure 1a) and Im9 bound to the DNase (Figure 1b) shows that the Im9 NH resonances are greatly perturbed by the DNase, with many peaks exhibiting complex-induced shifts of greater than 0.2 p.p.m. (^{15}N) and 0.04 p.p.m. (^1H). All the NH cross-peaks are broadened on formation of the complex owing to slower tumbling in solution; the broadening seems to be uniform, implying that there is no local conformational exchange within the complex. This observation is in agreement with NH exchange data. NH groups in free Im9 exchange within 3 days at pH 6.2 and 22 °C, but in Im9 bound to the E9 DNase NH exchange is still incomplete after 40 days, with a pattern of exchange that is broadly similar to that for the free protein.

Resonance assignments for the peptide NH groups of free Im9 (Figure 1a) have been obtained by conventional 2D and 3D NMR methods and have been reported elsewhere [20,21]. Attempts to assign Im9 resonances in the complex were greatly complicated because a simple comparison of the bound (Figure 1b) and free (Figure 1a) Im9 HSQC spectra is not sufficient to obtain firm assignments for all cross-peaks. This is particularly evident for the peak marked with a star on Figure 1(b), which has no obvious counterpart in Figure 1(a). The low dissociation rate for the complex means that it is not possible to obtain assignments by a titration method, and scalar coupling experiments (e.g. 3D ^{15}N -edited TOCSY–HMQC spectra) essential to the sequential assignment process gave spectra with insufficient cross-peaks, even with large numbers of scans to improve signal-to-noise, probably as a consequence of the short T_2 relaxation times of the slowly tumbling 24.5 kDa complex. Future NMR experiments on this complex will need to employ triple-resonance experiments with $^{13}\text{C}/^{15}\text{N}$ -labelled Im9 bound to unlabelled DNase [31], perhaps in concert with random deuteration of the Im9 [32,33], so that scalar coupling connectivities can be mapped by using relatively large heteronuclear coupling constants. This

Figure 1 600 MHz ^1H – ^{15}N HSQC spectra of free ^{15}N -labelled Im9 (a) and ^{15}N -labelled Im9 bound to the unlabelled DNase domain of ColE9 (b)

Both spectra were recorded at pH 6.2, 22 °C. Assignments are indicated on the spectra (see text for details).

approach should also allow the involvement of amino acid side chains to be investigated directly by NMR, something that the ^1H – ^{15}N data presented in this paper do not allow.

NMR resonance assignments

Without information from scalar coupling experiments, conventional assignment procedures are not applicable [31,34,35]. Assignment of the signals for the bound Im9 given in Figure 1(b) therefore relied heavily on the ^{15}N -edited 3D NOESY spectrum, which is illustrated in Figure 2 for sequential connectivities involving resonances of residues 48–55, as described below.

Regions of the sequence that exhibited minimal chemical shift perturbation on complexation (residues 3–19 and 76–86) were initially assigned by comparison with free Im9 and confirmed by the observation of sequential nuclear Overhauser enhancement (NOE) connectivities in the ^{15}N -edited 3D NOESY spectrum. The remaining signals were then assigned by first identifying candidate assignments for a resonance, based on chemical shifts in the free protein, and then using these as starting points for identifying sequential NOE connectivities (mainly NH–NH). This approach is illustrated in Figure 2, which shows sequential connectivities involving resonances of residues 48–55, including some of those most shifted by the binding of Im9 to the DNase (Figure 3). Comparison of the free and complexed ^1H – ^{15}N HSQC spectra (Figures 1a and 1b) shows that there are many possible assignments for residues 49 and 51–55 in the spectrum of the bound Im9. However, signals in the complex spectrum belonging to residues 48 and 50 can readily be identified by comparison with the spectrum of free Im9. Observation of sequential NOE connectivities (mainly NH–NH_{*i,i+1*}) were then used to assign the remaining portion of this stretch of the protein unambiguously. With this approach approx. 80 % of the peptide NH groups were assigned. In cases where no NH–NH NOEs were observed (e.g. residues 58–62), assignment was helped by detection of sequential NH–side chain NOEs. An additional nine resonances were assigned to groups of overlapping signals [e.g. residues 71 and 74 at approx. 8.6 (^1H) p.p.m. and 101 (^{15}N) p.p.m.]. By exclusion, most of the remaining signals must belong to residues 24–31, which were difficult to assign owing to the lack of observable NH–NH_{*i,i+1*} NOE connectivities allied with large complex-induced shifts. Residues 27–29, however, were assigned on the basis of detection of $\alpha\text{N}_{i,i+1}$ NOEs and comparison of their side chain shifts in the free spectrum. In total approx. 90 % of the backbone peptide resonances of Im9 complexed with ColE9 were assigned, although the assignments remain tentative without scalar coupling information.

The signal denoted by a star in Figure 1(b), with ^1H and ^{15}N chemical shifts of 11.67 and 88.0 p.p.m. respectively, displayed strong NOE cross-peaks to resonances with ^1H shifts at 5.98 and 7.62 p.p.m. No reciprocal NOEs were observed at 11.67 p.p.m. in the 3D ^1H – ^{15}N HMQC–NOESY spectrum, indicating that the signals at 5.98 and 7.62 p.p.m. originated from protons not directly attached to a ^{15}N nucleus; i.e. they belong to either aromatic protons or amide protons in the DNase protein. If it is assumed that these cross-peaks are intramolecular, then the most probable assignment of the peak at 5.98 p.p.m., on the basis of the spectrum of free Im9, is the H ^{β} proton of His-46, which is at 6.18 p.p.m. in free Im9. Additionally, the H ^{ϵ} resonance of His-46 in free Im9 is at 7.60 p.p.m.. In free Im9 the H ^{ϵ} and H ^{β} signals show NOEs to the amide protons of Ser-48 and Trp-74 respectively, and similar connectivities were seen to the peaks at 5.98 and 7.62 p.p.m. in the ^{15}N -edited 3D NOESY spectrum of the complex. Therefore the peak at 11.67 p.p.m. in Figure 1(b) (denoted with a star) is assigned to one of the NH ring protons of His-46.

We realize that the assignment strategy employed could lead to mis-assignment of resonances in the complex. However, the assignment strategy is unlikely to mis-assign contiguous stretches of NOEs, especially those involving residues in helices, and thus, because two of the most perturbed regions are helices II and III, we believe that the conservative approach we have adopted below for interpreting the complexation-induced shifts is valid. Also, the agreement between the NMR analysis and the results of the protein engineering experiments given below strengthens confidence in our identification of critical binding residues.

Figure 2 ^1H – ^1H strips taken from the 3D ^1H – ^{15}N HMQC-NOESY spectrum [26] of ^{15}N -labelled Im9 bound to unlabelled E9 DNase at pH 6.2 and 22 °C

NOESY mixing times used ranged from 80 to 150 ms; the spectrum illustrated had a mixing time of 100 ms. Slices were taken at ^{15}N frequencies corresponding to the backbone amide resonances of residues 48–55. Diagonal and sequential amide proton connectivities are shown. Artifacts present in the spectrum obscure information for the ^1H – ^1H slice taken at the ^{15}N frequency of Leu-52.

Figure 3 Graph showing the chemical shift perturbations of Im9 peptide ^1H and ^{15}N amide resonances caused by complexation of Im9 with E9 DNase

The complexation-induced chemical shift perturbations for residues indicated by stars could not be clearly determined either because of the lack of assignment or resolution in spectra of the bound Im9. The lightly shaded bars are the ^{15}N chemical shift changes and the dark bars are the ^1H chemical shift changes.

Figure 4 Surface and backbone worm representations of the NMR-derived solution structure of Im9 and the E9 DNase-binding region

All figures were created with the GRASP program [46]. Im9 contains four helices: I, residues 12–23; II, 30–44; III, 51–55; IV, 65–77. The colour coding for (a)–(d) reflects the chemical shift perturbations induced by binding to the E9 DNase (see text and Figure 3): red, residues whose NH group shows large chemical shift perturbations (1–3 and/or 0.2–0.6 p.p.m. for ^{15}N and ^1H respectively); blue, residues whose NH group shows intermediate chemical shift perturbations (0.2–1.0 and 0.04–0.20 p.p.m. for ^{15}N and ^1H respectively); white, residue amides that were either not perturbed or fell below these cut-offs. (a) Front face of helices II and III with several residues identified. (b) Top view of the molecule relative to (a). (c) Back view of the molecule relative to (a). (d) Bottom view of the molecule relative to (a). (e) Electrostatic surface potential of the surface viewed as in (a) is shown on the left. The surface potential was generated with the DELPHI program [45], with default size and charge parameters, inner and outer dielectric constants of 4 and 80 respectively, and an ionic strength of 0.2 M. In this figure red and blue indicate negative and positive potentials respectively. The worm figure on the right shows the NMR perturbations as in (a). (f) A comparison of conserved (white) with non-conserved (green) regions of the DNase-binding surface is shown on the left; the NMR perturbations as in panel (a) are shown on the right.

Assessment of complexation-induced chemical shift perturbations

A number of workers have used complexation-induced shifts in ^1H – ^{15}N HSQC spectra to map intermolecular interaction regions of protein–protein complexes. For example, Chen et al. [36] characterized the interaction between members of a bacterial phosphotransferase system, Jones et al. [37] identified the barnase binding site of barnase, and Clubb et al. [38] the binding surface of interleukin-8 complexed with the N-terminal fragment of its human receptor, whereas Gronenborn and Clore [39], Lian et al. [40] and Kato et al. [41] identified the contact surface of a streptococcal G-protein domain binding to a human Fc fragment. In all cases, unaffected resonances were ascribed to groups not involved in intermolecular contacts, and cut-offs of more than 0.2 p.p.m. and more than 0.02 (or more than 0.04) p.p.m. for ^{15}N

and ^1H chemical shift changes respectively were used to distinguish between affected and unaffected resonances.

Figure 3 shows the difference in ^1H and ^{15}N chemical shifts of Im9 on the formation of the complex with the E9 DNase. It is notable that approx. 75% of the peptide amide resonances of Im9 were affected to some degree by binding to the DNase. This amounts to approx. 60 amino acid residues distributed throughout the protein, considerably more than the number involved in protein–protein contacts in other complexes [42]. Thus, because it is unlikely that this number of residues are directly involved in binding the DNase, it is necessary to consider the origins of the individual chemical shift perturbations to delineate the interaction region. Figure 3 indicates that the perturbations can be considered in terms of small, intermediate and large changes in chemical shift, with the large category including amides shifted

by 1.0–3.0 p.p.m. and/or 0.2–0.6 p.p.m. in the ^{15}N and ^1H dimensions respectively, the intermediate category containing those resonances shifted by 0.2–1.0 p.p.m. and 0.04–0.2 p.p.m. in the ^{15}N and ^1H dimensions respectively, and the small category containing the remaining perturbed residues whose resonance shifted by at least 0.1 p.p.m. in one or both dimensions. Thus our large and intermediate categories are analogous to the cut-offs applied by other workers to determine interprotein contact residues. With these criteria 26 residues are highlighted: 11 in the most affected category (residues 27–29, 34, 35, 38, 49, 50 and 53–55) and 15 in the intermediate category (residues 23, 33, 36, 41–44, 46, 48, 52, 60, 63, 66, 68 and 69). The positions of these residues in the 3D structure of Im9 are shown in Figures 4(a)–4(d), with residues coloured red being the most affected and those coloured blue being in the intermediate category. In addition to these, 13 additional peptide amide resonances fall into the small category (residues 17, 19, 22, 39, 45, 51, 58, 59, 65, 72, 76, 78 and 79) but for clarity these are not highlighted in Figure 4.

To interpret the complex induced shifts, great care must be taken because Spitzfaden et al. [43] have shown that although perturbed resonances give good initial estimates of a binding surface, the effects are extended beyond the direct contact area, which in their case was described by NOE data. This is because chemical shifts are highly sensitive to the local magnetic environments of nuclei and these can be altered when two proteins form a complex for the following reasons.

1. There are changes in local environments because the NH is located at the interprotein binding region on the surface of one of the partners.

2. There are changes in chemical environments of NH groups not located at, but close to, the interprotein interface region. We call these induced shifts secondary shifts, and expect many perturbations to arise from this mechanism because a large number of the backbone NH groups are not exposed to the solvent. Note that only a minor conformational change, for example, of an aromatic ring, or the donor and acceptor contributing to a hydrogen bond, can lead to a substantial shift.

3. There are changes in chemical environments of NH groups as a consequence of a major conformational change, which is not restricted to the interface region, induced by complex formation.

Clearly, an observed perturbation does not necessarily indicate involvement at the complex interface. In many of the studies monitoring perturbations of ^1H – ^{15}N HSQC signals cited above, shifts for resonances belonging to nuclei far removed from the proposed binding area were observed. In these studies, however, major conformational changes on formation of a complex were known not to occur from crystallographic results.

In the absence of a 3D structure for the E9 DNase–Im9 complex we cannot discount the possibility that a conformational change accounts for some of the chemical shift perturbations in Im9. Indeed, stopped-flow fluorescence measurements show that association of the E9 DNase with Im9 is not only diffusion-controlled but goes through a conformational change before formation of the final stable complex [17]. The question arises, however, whether this change in conformation represents an alteration in the structure of Im9, the DNase or both, or merely a realignment of the two partners at the complex interface. Two pieces of data suggest that any conformational change occurs in the E9 DNase domain and not in the Im9: (1) the biosynthetic incorporation of 5-hydroxytryptophan, a sensitive fluorophore that can be selectively excited at 315 nm, into Im9 or the E9 DNase shows that only when the DNase carries the label can the conformational change be identified after excitation at 315 nm [17]; (2) consistent with the fluorescence data, the amide res-

onance of Trp-74, the sole tryptophan residue in Im9, and its surrounding residues remains relatively unperturbed on binding (Figure 3).

Taken together, the fluorescence and NMR data demonstrate that the region around Trp-74 of Im9, which is packed against residues from the loop connecting helix II and helix III, does not experience significant changes in microenvironment on binding to the DNase. This is an important point because in the following section we maintain that these two helices constitute the E9 DNase-binding site, a site that current data suggest is unlikely to go through a substantial change in conformation on binding the DNase.

In summary, then, previous NMR work on other interprotein complexes, together with the elimination of a significant complexation-induced conformational change in Im9, suggests that the residues whose NH groups are in the large and intermediate shift categories defined above (coloured red and blue in Figure 4) encompass the DNase-binding region on Im9.

Helix II and helix III contain the E9 DNase-binding region

Previous homologous recombination studies have suggested that residues 16–43 encompass the specificity-determining region of the immunity protein family [16], and from the recently determined structure of Im9 [20,21] this includes much of helix I and all of helix II. Further mutagenesis experiments by Wallis et al. [44] implicated residues 30–43 as being important for determining specificity, in particular Val-34, whose side chain is exposed to the solvent in the structure of free Im9.

Of the 26 perturbed amide resonances in the major and intermediate categories (Figure 3), 20 are from residues 27–55, which includes helix II, helix III and the loop that connects them, as well as the short loop connecting helix I with helix II. The relative dispositions of these residues in the structure of Im9 are shown in Figures 4(a)–4(d). Four out of five of the most highly perturbed surface residues (red) are located on the two helices, with another located in the loop linking helices I and II. These data lead us to suggest that part or all of the E9 DNase-binding site on the Im9 protein involves residues displayed from helices II and III. Indeed, the most perturbed residues on these helices form a nearly contiguous surface (see below and Figure 4a). Viewing the molecule from several angles reinforces the idea that the surface comprising helices II and III is likely to represent the E9 DNase-binding site (Figures 4a–4d).

The region spanning residues 27–29 comprises a solvent-exposed loop in the solution structure of Im9 that is relatively ill-defined [21]. This region also shows large complexation-induced chemical shift changes (Figure 3) and thus might be directly involved at the complex interface or structurally perturbed on complexation.

The perturbations to the chemical shifts of the NH of the exposed hydrophobic residue Val-34 of helix II are among the highest seen in Im9 (–0.51 p.p.m. and –1.2 p.p.m. for ^1H and ^{15}N respectively), and there are also large perturbations for residues surrounding Val-34, e.g. Leu-33, Lys-35, Leu-36 and Thr-38. These residues are all part of helix II, which remains a helix in the complex on the basis of the observation of sequential amide ^1H connectivities. Because the observed shifts throughout this helix seem too large to be accounted for by alterations in H-bond strengths alone, we conclude that at least part of this helix is at the complex interface, which is consistent with the available biochemical data. We also note from the solution structure of Im9 that the side chains of residues 30, 31, 35, 38 and 42, in addition to 34, are on the solvent-exposed face of helix II, and

suggest that at least some of these residues are important for conferring specificity.

The short helix III in Im9 forms a contiguous surface with helix II and, as observed by Osborne et al. [21], contains hydrophobic residues on the surface (Tyr-54 and Tyr-55) that are in close proximity to the hydrophobic Val-34 on the surface of helix II. Residues on this helix and on the turn joining helix II and helix III are affected considerably on complexation (Figures 4a–4d). Interestingly, a large number of residues on helix III and the 40s turn are conserved throughout the DNase immunity proteins and consequently this region has not been previously implicated in binding. We return to this point later.

In contrast with the substantial perturbations seen in helix II as a result of binding to the E9 DNase, only one amide group (Cys-23) from helix I falls within either of the cut-offs used to assign the binding surface (Figures 3 and 4a). Several of the amides of helix I do, however, show perturbations in a single dimension (see above) within the lower cut-off. Most of these (residues 15, 17, 19 and 22) are mainly buried within the protein [21] and their side chains form contacts with helix II and helix III. The exposed surface of helix I, comprising residues 12, 14, 16 and 20, is not perturbed by DNase binding (Figure 4c). Chemical modification experiments have previously shown that Cys-23 of helix I becomes inaccessible to modification on binding ColE9, suggesting that it is part of the DNase-binding site [44]. However, further experiments demonstrated that when Cys-23 is modified with a variety of alkylating reagents of differing chain lengths the derivatized Im9 proteins can still inhibit the DNase activity of ColE9 but only up to a certain derivative size, implying that this residue is located near, but not at, the DNase-binding site [44]. The position of Cys-23 relative to the surface of helices II and III is consistent with these data (Figures 4a and 4d). Thus we conclude that helix I is not directly involved at the complex interface. From the structural and NMR data we can also discount residues of helix IV, which spans residues 65–76, from being involved at the interface. Although several amide resonances at the N-terminal end of this helix show intermediate shifts, most of these are buried residues (e.g. Val-68).

The collision between Im9 and the E9 DNase is steered electrostatically and occurs at essentially the rate of diffusion [17]. The implication from the strongly salt-dependent association kinetics is that the E9 DNase-binding surface of Im9 should be negatively charged. The electrostatic potential of the putative DNase-binding surface of Im9 has been calculated with the DELPHI program [45] and is depicted in Figure 4(e). It is clear from this representation that the DNase-binding surface of Im9 is highly negatively charged; almost a third of the residues in this region are acidic and this accounts for half the total number of acidic residues in Im9. No other surface region of Im9 shows such a high concentration of negative charge. The highest density of negative charge originates from the side chain carboxylates of residues 41, 42 and 51, although the exposed backbone carbonyl groups of residues 46 and 48 and the hydroxy groups of Tyr-54 and Tyr-55 also contribute to the negative charge density.

Protein engineering experiments

The strategy we have adopted to delineate the E9 DNase-binding site on Im9 by amide chemical shift perturbations has identified a binding surface that is consistent with the available biochemical and genetic data on this complex. Nevertheless the absence of scalar coupling information, the categorization of the complexation-induced chemical shift perturbations and the absence of direct information on the involvement of residue side chains raises questions about the reliability of this strategy. To

address these issues further, and thereby validate our NMR approach to delineating the binding surface, we replaced five residues with alanine, expressed and purified the mutant proteins and determined the dissociation rate constants for the mutant Im9–E9 DNase complexes. The residues selected for replacement are all on the surface of Im9 and represent the different categories defined by the complexation-induced chemical shift changes: ‘red’ for most perturbed, ‘blue’ for intermediate and ‘white’ for the rest (Figures 3 and 4). The dissociation rate constants were chosen as kinetic indicators of the binding surface on Im9 because previous analyses of cognate and non-cognate immunity protein–E9 DNase complexes have shown that although their association rate constants are almost identical their dissociation rate constants vary considerably [17,18]. This is consistent with other studies of protein complexes, which show that in general protein complex off-rates determine the affinity of the complex [42]. Rate constants for dissociation were obtained by subunit exchange experiments in which wild-type Im9, which had been biosynthetically radiolabelled with [³H]leucine, was exchanged into unlabelled mutant complexes (Figure 5) [17]. All the variant proteins studied had similar fluorescence spectra to that of wild-type Im9, and behaved similarly to it on chromatographic columns, which is consistent with the alanine substitutions not producing major structural perturbations.

The dissociation rate constant (k_{off}) for wild-type Im9 at pH 7.0 and in 50 mM Mops buffer containing 200 mM NaCl and 1 mM dithiothreitol was $1.5 \times 10^{-6} \text{ s}^{-1}$ ($t_{1/2} = 128 \text{ h}$; Figure 5, inset) which compares favourably with our previous estimate of $2.2 \times 10^{-6} \text{ s}^{-1}$ (determined by chasing out radioactive Im9 from the wild-type complex with unlabelled Im9 [17]). A single ‘white’ residue on the putative DNase-binding surface was analysed, Glu³¹ → Ala, which showed a similar dissociation rate constant ($k_{\text{off}} = 2.4 \times 10^{-6} \text{ s}^{-1}$; $t_{1/2} = 80 \text{ h}$) to that of wild-type Im9, which is consistent with this residue being unimportant for binding. Two ‘blue’ residues were also mutated to alanine (Cys-23 and Glu-42); both changes affected the dissociation rate but only by 3–4-fold ($k_{\text{off}} = 6.1 \times 10^{-6}$ and $4.6 \times 10^{-6} \text{ s}^{-1}$; $t_{1/2} = 31$ and 42 h respectively). Conversely, the two ‘red’ residues that were chosen for mutation had a marked effect on the dissociation rate constant; k_{off} for Val³⁴ → Ala (helix II) was $2.5 \times 10^{-4} \text{ s}^{-1}$ ($t_{1/2} = 0.8 \text{ h}$), whereas k_{off} for Tyr⁵⁴ → Ala (helix III) was $2.8 \times 10^{-3} \text{ s}^{-1}$ ($t_{1/2} = 0.07 \text{ h}$) (Figure 5). Thus the relative effects of the alanine substitutions are fully in accord with the chemical shift categories used to define the E9 DNase-binding surface.

Common contact sites and immunity protein–DNase specificity

The difference in binding affinity between cognate and non-cognate E9 DNase–immunity protein complexes is 12 orders of magnitude, yet the binding of each immunity protein still results in the inhibition of DNase activity. The affinity differences between the immunity proteins for the E9 DNase are governed solely by the dissociation rate constant for the complex because the association rate constants are almost identical for cognate and non-cognate complexes [16]. These observations suggest that a common contact site exists on each inhibitor protein, possibly one that is involved in inactivating the nuclease. If such a site exists, then to accommodate the substantial differences in binding affinity between the inhibitor proteins the site must either contribute very little to the binding energy, and/or its binding interactions must be modulated by surrounding residues. This is consistent with our previous proposal that a substantial part of the high binding affinity of the E9 DNase–Im9 complex results from just a small number of short-range interactions within a relatively large binding surface [16]. The decreased affinity for

Figure 5 Radioactive exchange data for wild-type and mutant Im9 proteins

For the rapidly exchanging complexes of Tyr⁵⁴ → Ala or Val³⁴ → Ala Im9 complexed with the E9 DNase, first-order rates of dissociation were obtained by fitting the data (plotted as the percentage loss of bound immunity) to single-exponential rate equations, as shown in the main figure for Tyr⁵⁴ → Ala. For the slowly exchanging complexes (which were not taken to completion) including wild-type Im9, Glu³¹ → Ala, Glu⁴² → Ala and Cys²³ → Ala, first-order rates of exchange were estimated from semi-logarithmic plots (inset). Dissociation rate constants for these complexes are given in the text. For clarity, the data for Cys²³ → Ala and Val³⁴ → Ala Im9 are not shown in the Figure.

non-cognate complexes could then arise from a small number of unfavourable short-range interactions negating favourable ones. One way to construct a common binding site with different specificities is to base it on both highly conserved and variable regions of the protein. The sequence composition of the putative E9 DNase-binding site on Im9 matches that expected for such a surface. Figure 4(f) shows that half the binding surface, comprising most of the residues in helix III, is highly conserved within the immunity protein family, whereas the other half, comprising residues displayed from helix II, is highly variable in sequence. A particularly striking feature of this surface composition is the close proximity of the two invariant tyrosines at the C-terminal end of helix III, Tyr-54 and Tyr-55, with Val-34 (a specificity-determining residue) forming a hydrophobic patch on the surface of Im9. In other immunity proteins position 34 is either an aspartic acid (Im8 and Im7) or an asparagine (Im2) residue. Thus the relative contributions to the interaction binding energies of E9 DNase-immunity protein complexes made by these residues might be the key to determining the nature of the specificity in these complexes. Already it can be seen from the mutagenesis data in this paper that two of these residues (Val-34 and Tyr-54) have a substantial effect on binding when replaced with alanine.

It is interesting to note that the greatest concentration of negative charge on the proposed binding surface originates from helix III and the turn attaching this helix to helix II. These residues are mostly conserved throughout the immunity protein family (Figures 4e and 4f) and thus this could account for the similar values of k_{on} observed for the colicin E9 DNase with all the DNase immunity proteins and the finding that all these associations are highly salt-dependent [16,17]. Additional mutagenesis experiments with Im9 and structural studies of the DNase will be required to characterize this further.

The binding model we have presented centres on helices II and III of Im9 constituting the binding site for the E9 DNase. This

is at variance with the proposal of Chak et al. [22] that helices I and II (both highly variable in the immunity protein family) are the main components of the nuclease-binding site for an immunity protein. However, we note that Chak et al. did not present any additional data to support their proposal but instead relied heavily on previous genetic evidence and sequence alignment data [22]. The structures of Im9 and Im7 are very similar and because both can inhibit the activity of the E9 DNase it is to be expected that a similar surface is used by both proteins to inactivate the nuclease, albeit with markedly differing affinities [16]. The heteronuclear NMR experiments and the results with the alanine mutants reported in this paper suggest that this surface is composed of variant and invariant residues displayed from helices II and III respectively. This model is fully consistent with previous genetic data from our laboratory [44], which suggested that helix II is important in defining the nuclease specificity of an immunity protein, and suggests for the first time that adjoining conserved residues from helix III also participate in binding to the endonuclease.

We thank Ann Reilly for her technical assistance, Ansgar Pommer for providing the E9 DNase for many of the experiments described in the paper, and the Biotechnology and Biological Sciences Research Council and the Engineering and Physical Sciences Research Council for supporting the work of the UEA Colicin Research Group via their Biomolecular Sciences Committee. G. R. M. thanks the Wellcome Trust for its award of a Research Leave Fellowship.

REFERENCES

- Schleif, R. (1988) *Science* **241**, 1182–1187
- Saenger, W. and Heinemann, U. (1989) *Protein–Nucleic Acid Interactions*, Macmillan Press, New York
- Sauer, R. T. (1992) *Methods Enzymol.* **208**, 700
- Davies, D. R. and Padlan, E. A. (1990) *Annu. Rev. Biochem.* **59**, 439–473
- Kelley, R. F. and O'Connell, M. P. (1993) *Biochemistry* **32**, 6828–6835
- Hawkins, R. E., Russell, S. J., Baier, M. and Winter, G. (1993) *J. Mol. Biol.* **234**, 958–964
- Lee, F. S., Shapiro, R. and Vallee, B. L. (1989) *Biochemistry* **28**, 225–230
- Schreiber, G. and Fersht, A. R. (1993) *Biochemistry* **32**, 5145–5150
- Guillet, V., Laphorn, A., Hartley, R. W. and Manguen, Y. (1993) *Structure* **1**, 165–176
- Buckle, A. M., Schreiber, G. and Fersht, A. R. (1994) *Biochemistry* **33**, 8878–8889
- Cunningham, B. C. and Wells, J. A. (1989) *Science* **244**, 1081–1085
- De Vos, A. M., Ultsch, M. and Kossiakoff, A. A. (1992) *Science* **255**, 306–312
- Clackson, T. and Wells, J. A. (1995) *Science* **267**, 383–386
- Bode, W. and Huber, R. (1992) *Eur. J. Biochem.* **204**, 433–451
- James, R., Kleanthous, C. and Moore, G. R. (1996) *Microbiology* **142**, 1569–1580
- Wallis, R., Leung, K.-Y., Pommer, A. J., Videler, H., Moore, G. R., James, R. and Kleanthous, C. (1995) *Biochemistry* **34**, 13751–13759
- Wallis, R., Moore, G. R., James, R. and Kleanthous, C. (1995) *Biochemistry* **34**, 13743–13750
- Wallis, R., Reilly, A., Barnes, K., Abell, C., Campbell, D. G., Moore, G. R., James, R. and Kleanthous, C. (1994) *Eur. J. Biochem.* **220**, 447–454
- Wallis, R., Reilly, A., Rowe, A., Moore, G. R., James, R. and Kleanthous, C. (1992) *Eur. J. Biochem.* **207**, 687–695
- Osborne, M. J., Lian, L.-Y., Wallis, R., Reilly, A., James, R., Kleanthous, C. and Moore, G. R. (1994) *Biochemistry* **33**, 12347–12355
- Osborne, M. J., Breeze, A., Lian, L.-Y., Reilly, A., James, R., Kleanthous, C. and Moore, G. R. (1996) *Biochemistry* **35**, 9505–9512
- Chak, K.-F., Safo, M. K., Ku, W.-Y., Hsieh, S.-Y. and Yuan, H. S. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6437–6442
- Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113**, 967–974
- Braunschweiler, L. and Ernst, R. R. (1983) *J. Magn. Reson.* **53**, 521–528
- Shaka, A. J., Barker, P. B. and Freeman, R. (1985) *J. Magn. Reson.* **64**, 547–552
- Zuiderweg, E. R. P. and Fesik, S. W. (1989) *Biochemistry* **28**, 2387–2391
- Bax, A., Davis, D. G. and Sarkar, S. K. (1985) *J. Magn. Reson.* **63**, 230–234
- Norwood, T. J., Boyd, J., Heritage, J. E., Softe, N. and Campbell, I. D. (1990) *J. Magn. Reson.* **87**, 488–501
- Bax, A., Ikura, J., Kay, L. E., Torchia, D. A. and Tschudin, R. (1990) *J. Magn. Reson.* **86**, 304–318
- Sarkar, G. and Sommers, S. S. (1990) *BioTechniques* **8**, 404–407
- Bax, A. and Grzesiek, S. (1993) *Acc. Chem. Res.* **26**, 131–138

-
- 32 Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R. and Kay, L. E. (1994) *J. Am. Chem. Soc.* **116**, 11655–11666
- 33 Nietlispach, D., Clowes, R. T., Broadhurst, R. W., Ito, Y., Keeler, J., Kelly, M., Ashurst, J., Oschkinat, H., Dommelle, P. J. and Laue, E. D. (1996) *J. Am. Chem. Soc.* **118**, 407–415
- 34 Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York
- 35 Englander, S. W. and Wand, A. J. (1987) *Biochemistry* **26**, 5954–5958
- 36 Chen, Y., Reizer, J., Saier, Jr., M. H., Fairbrother, W. J. and Wright, P. E. (1993) *Biochemistry* **32**, 32–37
- 37 Jones, D. N., Bycroft, M., Lubienski, M. J. and Fersht, A. R. (1993) *FEBS Lett.* **331**, 165–172
- 38 Clubb, R. T., Omichinski, J. G., Clore, G. M. and Gronenborn, A. M. (1994) *FEBS Lett.* **338**, 93–97
- 39 Gronenborn, A. M. and Clore, G. M. (1993) *J. Mol. Biol.* **233**, 331–335
- 40 Lian, L.-Y., Barsukov, I. L., Derrick, J. P. and Roberts, G. C. K. (1994) *Struct. Biol.* **1**, 355–357
- 41 Kato, K., Lian, L.-Y., Barsukov, I. L., Derrick, J. P., Kim, H., Tanaka, R., Yosino, A., Siraishi, M., Shimada, I., Arata, Y. and Roberts, G. C. K. (1995) *Structure* **3**, 79–85
- 42 Janin, J. and Chothia, C. (1990) *J. Biol. Chem.* **265**, 16027–16030
- 43 Spitzfaden, C., Weber, H.-P., Braun, W., Kallen, J., Wider, G., Widmer, H., Walkinshaw, M. D. and Wüthrich, K. (1992) *FEBS Lett.* **300**, 291–300
- 44 Wallis, R., Moore, G. R., Kleanthous, C. and James, R. (1992) *Eur. J. Biochem.* **210**, 923–930
- 45 Nicholls, A. and Honig, B. (1991) *J. Comp. Chem.* **12**, 435–445
- 46 Nicholls, A., Sharp, K. and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11**, 281–296
-

Received 1 October 1996/16 December 1996; accepted 6 January 1997