Journal of Biomolecular NMR **29:** 395–398, 2004.

Letter to the Editor: NMR structure of a variant 434 repressor DNA-binding domain devoid of hydroxyl groups

Hideo Iwai[a,b], Gerhard Wider[a,*] & Kurt Wüthrich[a]

[a]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Hönggerberg, CH-8093 Zürich, Switzerland;* [b]*Present address: Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, SK, Canada S7N 5C9*

Received 3 February 2004; Accepted 15 March 2004

Key words: NMR structure, protein hydration, 434 repressor

Abbreviations: 434(1–63) – N-terminal 63-residue DNA-binding domain of the phage 434 repressor; dh434(0–63) – variant 434 repressor DNA-binding domain devoid of hydroxyl groups; P22c2(1–76) – N-terminal 76-residue fragment of the phage P22 c2 repressor; SDS-PAGE – SDS polyacrylamide gel electrophoresis; COSY – correlation spectroscopy; TOCSY – total correlation spectroscopy; NOE – nuclear Overhauser effect; RMSD – root-mean-square deviation.

Biological context

The presently described cloning, expression and NMR structure determination of variant 434 repressor DNA-binding domains with residues 1–63, 434(1–63) (Mondragon et al., 1989; Neri et al., 1992; Pervushin et al., 1996), was initiated in the context of NMR studies on the location and dynamics of water molecules in the surface hydration shell of proteins (Otting et al., 1991; Wider, 1998). Since amino acid side chain hydroxyl groups interfere with such studies (Liepinsh et al., 1992) we prepared a 434(1–63) devoid of all serine and threonine residues. With the use of this newly designed protein we investigated the feasibility of identifying the complete layer of hydration waters on the surface of globular proteins. In addition, this new construct, dh434(0–63), unexpectedly turned out to have significantly increased thermal stability. Here, we describe a structure determination of dh434(0–63) and investigate the basis of this increased stability.

Methods and results

The plasmid for the expression of dh434(0–63) was derived from the expression vector for 434(1–63), pT7-7/434A. Because the N-terminal nonapeptide of 434(1–63) has a high content of Ser residues, it was genetically replaced with the corresponding sequence of P22c2(1–76) which has a very similar three-dimensional structure as 434(1–63) (Sevilla-Sierra et al., 1994). Further amino acid replacements needed to obtain dh434(0–63) (Figure 1) were introduced by gene synthesis with the assembly polymerase chain reaction. To improve the solubility of the resulting new construct, Ala at position 50 was replaced by Arg with site-directed mutagenesis. All the mutations were confirmed by DNA sequencing. The plasmid was then transfected into *E. coli* BL21(DE3) cells. The cells were grown at 37 °C in LB-medium until an OD_{600} of 0.6 was reached, and then induced with IPTG (fi-

```
                0          10          20          30
dh434(0-63) MLMGERIRARRIQLGLNQAELAQKVGVDQQA

434(1-63)   -SISSRVKSK                   TT  S

               31          40          50          60
            IEQLENGKAKRPRFLPELARALGVAVDWLLNGA

                     T           S     S        T
```

Figure 1. Comparison of the amino acid sequences of the newly designed protein dh434(0–63) and of its wild type analogue 434(1–63). dh434(0–63) is a hybrid containing the N-terminal decapeptide segment from the P22c2 repressor (in italics) and the segment 10–63 from 434(1–63) with all residues containing hydroxyl groups replaced by inert residues (see text). For 434(1–63) (lower line) only the variations relative to dh434(0–63) are indicated.

*To whom correspondence should be addressed. E-mail: gsw@mol.biol.ethz.ch

Table 1. Characterization of the 20 energy-refined DYANA conformers representing the NMR structure of dh434(0–63)[a]

DYANA residual target function value (Å^2)[b]	0.57 ± 0.15 (0.33...0.82)
AMBER energy (kcal/mol)	−2852 ± 22 (−2529...−2333)
NOE violations > 0.1 Å	0.20 ± 0.40 (0...1)
angle violations > 2.5°	0.35 ± 0.48 (0...1)
RMSD[c] (Å): N, C$^\alpha$, and C′(1–63)	0.39 ± 0.06
RMSD[c] (Å): all heavy atoms (1–63)	0.83 ± 0.06
RMSD[d] (Å): dh434 vs. 434(1–63)	0.62

[a]Average values ± standard deviations for the group of 20 DYANA conformers after energy minimization are given; the minimum and maximum values for the individual conformers are given in parentheses.
[b]Before energy minimization.
[c]Averages are given of the pairwise RMSD values between each of the 20 energy-refined DYANA conformers and the mean solution structure.
[d]Calculated for the mean positions of the backbone heavy atoms, N, C$^\alpha$, and C′ of residues 1 to 63 relative to the mean structure of 434(1–63) (Pervushin et al., 1996).

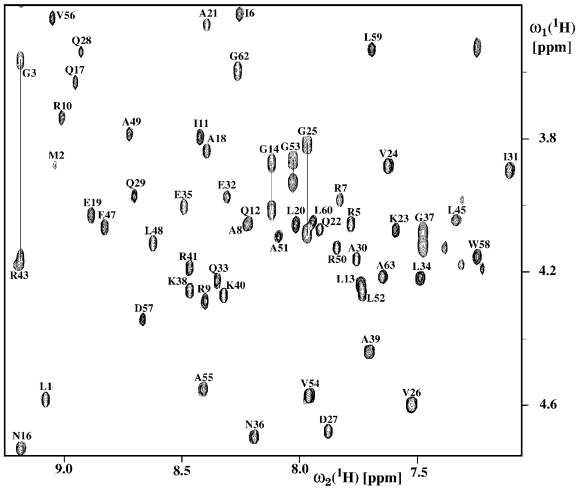

Figure 2. Fingerprint region of the 750 MHz phase-sensitive [^{1}H,^{1}H]-TOCSY spectrum of dh434(0–63) in 90% H_2O/10% D_2O. Resonance assignments are given by the one-letter amino acid code and the sequence position.

nal concentration 0.5 mM) for 6 h. The cells were harvested and lysed, and the extract was centrifuged at 35,000 × *g* for 1 h. The supernatant was applied to a SP-Sepharose column, which had been pre-equilibrated with a solution of 50 mM sodium acetate at pH 4.1, 1 mM EDTA, 1 mM β-mercaptoethanol, 1 mM Pefabloc SC and 10 mM NaCl. The dh434(0–63) was eluted by a linear gradient from 10 mM to 500 mM NaCl, and protein-containing fractions were analyzed by SDS-PAGE. All relevant fractions were pooled and dialyzed extensively against distilled water. The protein solution was diluted twice with the starting buffer and loaded onto a cellulose phosphate column (Whatman P11) which has been pre-equilibrated with a buffer of 50 mM potassium phosphate at pH 6.1, 1 mM β-mercaptoethanol and 30 mM NaCl. The protein was then eluted by a linear gradient from 30 mM to 500 mM NaCl in the same buffer. The fractions were analyzed by SDS-PAGE, and those containing the pure protein were pooled for further extensive dialysis against distilled water.

The complete amino acid sequence of the purified dh434(0–63) (Figure 1) was confirmed by Edman degradation. Dynamic light scattering data of a 1.8 mM solution of dh434(0–63) in 25 mM potassium phosphate at pH 5.3 indicated a mono-disperse protein preparation with an estimated molecular mass of 9 kDa. Surprisingly, dh434(0–63) was found to be significantly more stable than 434(1–63) against high temperature and chemical denaturants. While 434(1–63) has a denaturation temperature of T_m = 68 °C, T_m = 85 °C for dh434(0–63). The midpoints of the GdnHCl unfolding curves are at 1.7 and 2.1 M for 434(1–64) and dh434(0–63), respectively.

The [^{1}H,^{1}H]-TOCSY spectrum of Figure 2 illustrates the high quality of the NMR data obtained for dh434(0–63). NMR spectra were measured with a 5 mM solution at pH 4.8 of unlabeled dh434(0–63) in 90% H_2O/10% D_2O containing 25 mM potassium phosphate. All spectra were recorded at 750 MHz at 13 °C. Sequence-specific assignments were obtained using sequential NOEs and ^{1}H–^{1}H scalar couplings (Wüthrich, 1986). Vicinal spin–spin coupling constants $^3J_{HN\alpha}$ were measured in a [^{15}N,^{1}H]-COSY spectrum by inverse Fourier transformation (Szyperski et al., 1992) and $^3J_{\alpha\beta}$ data were extracted from a

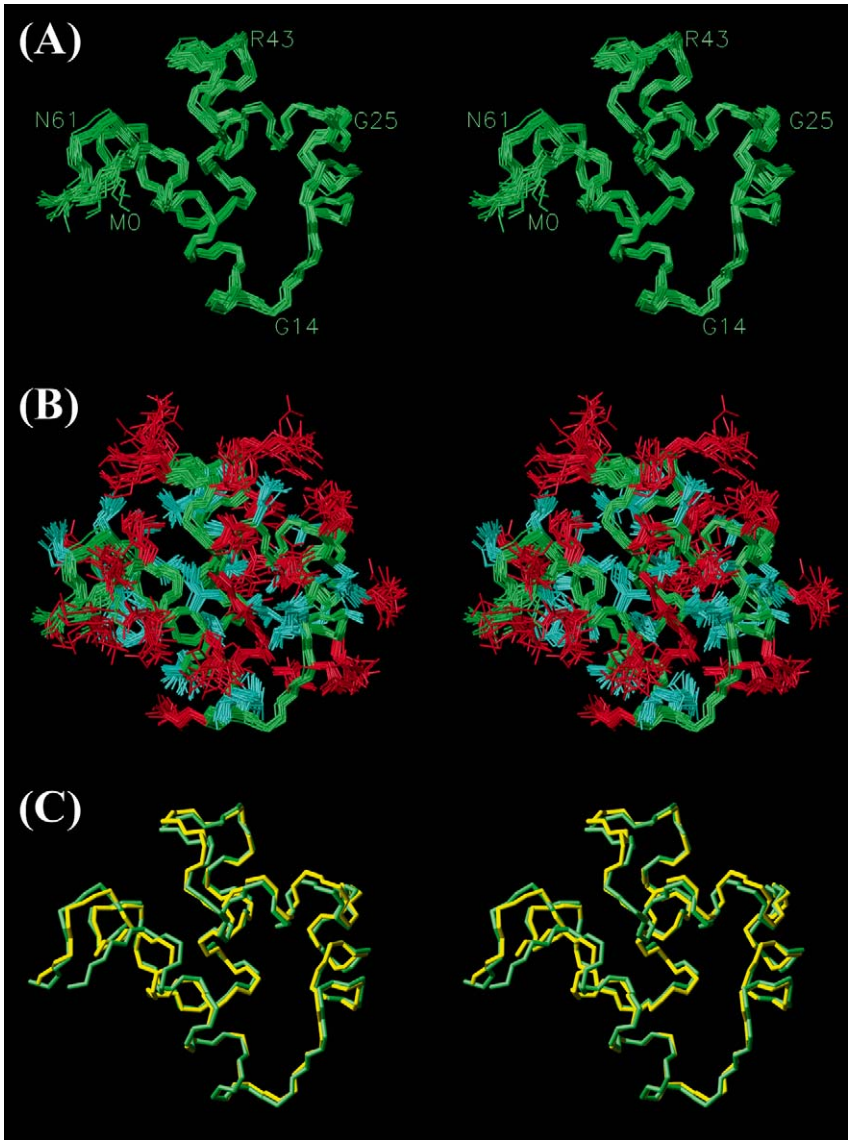

Figure 3. The NMR structure of dh434(0–63) shown with stereoviews. (A) Polypeptide backbone of 20 energy-minimized NMR conformers with superposition for best fit of the backbone atoms N, C^{α} and C'. (B) Same superposition as in (a) showing also the side-chain heavy atoms. The backbone, the 'best-defined' side-chains (i.e., side-chains for which the global displacements are smaller than 0.8 Å), and the remaining side-chain are colored in green, blue, and red, respectively. (C) Stereoview of the superposition for best fit of the backbone atoms N, C^{α} and C' of residues 1 to 63 of the mean structures of dh434(0–63) (green), 434(1–63) (yellow). This figure was prepared with the program MOLMOL (Koradi et al., 1995).

[^{1}H,^{1}H]-E. COSY spectrum (Griesinger et al., 1985). Amide proton exchange rates were measured using a lyophilized protein sample obtained from a 1 mM solution of ^{15}N-labeled dh434(0–63) in H_2O at pH 4.8 containing 25 mM potassium phosphate. The protein was redissolved in D_2O and a series of 2D [^{15}N, ^{1}H]-COSY spectra were recorded at 13 °C. The complete ^{1}H assignments have been deposited in the BioMagResBank (accession code BMRB-6084).

The NMR solution structure of dh434(0–63) was calculated from a total of 1026 NOE distance constraints, including 560 long-range, 94 medium-range, 78 sequential and 294 intra-residual NOEs. In addition, 135 dihedral angle constraints were obtained from scalar coupling constants and local NOEs using

the program HABAS (Güntert et al., 1991), which together with the program GLOMSA (Güntert et al., 1991) also provided a number of stereospecific assignments for β-methylene protons. For the input of the DYANA calculations no explicit constraints for hydrogen bonds were included (Güntert and Wüthrich, 1991; Güntert et al., 1997). The final calculation was started with 50 random structures and used a standard DYANA annealing protocol. The 20 best DYANA conformers were subjected to energy-minimization in a water bath using the program OPAL (Luginbühl et al., 1996) (Table 1).

The NMR structure (Figure 3A) reveals that dh434(0–63) consists of five helices with residues 2 to 13 (helix I), 17 to 24 (helix II), 28 to 36 (helix III), 45 to 52 (helix IV), and 56 to 61 (helix V). The molecular architecture of dh434(0–63) is essentially identical to that of 434(1–63) (Figure 3C). The RMSD of 0.6 Å between 434(1–63) and dh434(0–63) (Table 1) indicates that there is no significant deviation between the backbone conformations (Figure 3C), in spite of the 16 amino acid replacements which make the protein significantly more stable against denaturation. The amide proton exchange data of dh434(0–63) show slow exchange rates for the five helices and throughout the exchange rates for dh434(0–63) are slower than the corresponding data for 434(1–63), which is in line with the aforementioned increase of T_m from 68 °C for 434(1–63) to 85 °C for dh434(0–63) (Wagner and Wüthrich, 1978). The NMR structure of dh434(0–63) could be obtained with higher precision, which suggests a more compact structure than for 434(1–63). The atom coordinates have been deposited in the protein data bank (submission code PDB-ISQ8).

Discussion and conclusions

The newly designed protein dh434(0–63) has a nearly identical backbone conformation to that of the natural protein, 434(1–63). The most striking difference is the significantly higher stability of dh434(0–63) against thermal and chemical denaturation. In the three-dimensional structure of dh434(0–63) (Figure 3B), a larger number of both hydrophobic contacts and hydrogen bonds were observed when compared to 434(1–63). These features are reminiscent of observations made in the three- dimensional structures of thermophilic proteins, and these additional non-bonding contacts may be the main reason for the hyperstability of dh434(0–63) (Chan et al., 1995; Goldman, 1995; Korolev, 1995). The introduction of a number of additional hydrophobic side chains of valine and alanine into 434(1–63) (Figure 1) increases the internal van der Waals contact area, and overall the hydrophobic core of dh434(0–63) is thus more tightly packed. Among the additional hydrogen bonds, the side-chain of Arg9 (Lys9 in 434(1–63)) forms novel hydrogen bonds with the backbone oxygen atoms of either Leu52, Gly53 or Leu52 and Gly53 in individual ones of the 20 NMR conformers. These hydrogen bonds connect the two sub-domains of dh434(0–63) and may thus make a particularly large contribution to the observed increased protein stability.

The use of the structure of dh434(0–63) for a detailed characterization of the surface hydration which will be published elsewhere.

Acknowledgements

Financial support was obtained from the Schweizerischer Nationalfonds (project 31.49047.96). HI was in part supported by JSPS Research Fellowship for Young Scientists.

References

Chan, M.K., Mukund, S., Kletzin, A., Adams, M.W. and Rees, D.C. (1995) *Science*, **267**, 1463–1469.
Goldman, A. (1995) *Structure*, **3**, 1277–1279.
Griesinger, C., Sørensen, O.W. and Ernst, R.R. (1985) *J. Am. Chem. Soc.*, **107**, 6394–6396.
Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.
Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.
Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.
Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph.*, **14**, 51–55.
Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 9264–9268.
Liepinsh, E., Otting, G. and Wüthrich, K. (1992) *J. Biomol. NMR*, **2**, 447–465.
Luginbühl, P., Güntert, P., Billeter, M. and Wüthrich, K. (1996) *J. Biomol. NMR*, **8**, 136–146.
Mondragon, A., Subbiah, S., Almo, S.C., Drottar, M. and Harrison, S.C. (1989) *J. Mol. Biol.*, **205**, 189– 201.
Neri, D., Billeter, M. and Wüthrich, K. (1992) *J. Mol. Biol.*, **223**, 743–767.
Otting, G., Liepinsh, E. and Wüthrich, K. (1991) *Science*, **254**, 974–980.
Pervushin, K., Billeter, M., Siegal, G. and Wüthrich, K. (1996) *J. Mol. Biol.*, **264**, 1002–1012.
Sevilla-Sierra, P., Otting, G. and Wüthrich, K. (1994) *J. Mol. Biol.*, **235**, 1003–1020.
Szyperski, T., Güntert, P., Otting, G. and Wüthrich, K. (1992) *J. Magn. Reson.*, **99**, 552–560.
Wagner, G. and Wüthrich, K. (1978) *Nature*, **275**, 247–248.
Wider, G. (1998) *Progr. NMR Spectrosc.*, **32**, 193–275.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.