Assignment of the backbone $^{1}H,^{15}N,^{13}C$ NMR resonances and secondary structure of a double-stranded RNA binding domain from the *Drosophila* protein staufen

Mark Bycroft[a,*], Mark Proctor[a], Stefan M.V. Freund[a], Daniel St. Johnston[b]

[a]*Cambridge Centre for Protein Engineering, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK*
[b]*Wellcome/Cancer Research Campaign Institute, Tennis Court Road, Cambridge, CB2 1QR, UK*

Received 14 October 1994; revised version received 5 February 1995

Abstract NMR spectroscopy has been used to determine the secondary structure of one of the double-stranded RNA binding domains from the *Drosophila* protein staufen. The domain has an $\alpha\beta\beta\beta\alpha$ arrangement of secondary structure, with the β strands forming an antiparallel β sheet. The secondary structure differs from that found in the RNP RNA binding domain.

Key words: RNA binding domain; NMR; Secondary structure

1. Introduction

A variety of protein domains that interact with RNA have been identified [1]. Amongst these is a domain of approximately 70 amino acids that binds to double-stranded RNA [2]. This domain was first identified in the product of the *Drosophila* gene for staufen, dsRNA activated protein kinase, and the *Xenopus* protein xlrbpa. Subsequent database searching showed that the domain is present in several other proteins. Recently the dsRNA binding domain has been detected in a number of RNA helicases [3]. The domain occurs five times in the protein staufen. Staufen is a maternal effect gene that is required for the localisation of mRNA. One of the RNA binding domains of staufen (residues 575–647) has been shown to bind dsRNA in vitro [2]. In this paper we report the assignment of $^{1}H,^{15}N,^{13}C$ resonances in the NMR spectrum of this domain and the determination of its secondary structure.

2. Materials and methods

The cloning of the third dsRNA binding domain of the staufen gene into the expression vector pGEX 2T has been described previously [2]. Glutathione *S*-transferase RNA binding domain fusion protein was purified using glutathione-Sepharose (Pharmacia) according to the manufacturer's instructions. Purified fusion protein was cleaved using thrombin (Sigma) and the RNA binding domain was purified away from the carrier protein using a heparin-Ultrogel column (IBF). Uniformly ^{15}N and $^{15}N,^{13}C$ labelled protein was prepared from cells grown in M9 minimal media containing ^{15}N ammonium chloride and ^{13}C glycerol (Isotec).

NMR samples were approximately 2 mM protein in 10 mM potassium phosphate, pH 6.8, 3 mM deuterated DDT (Isotec). NMR spectra were recorded on a Bruker AMX 500 spectrometer equipped with a $^{1}H,^{15}N,^{13}C$ triple resonance probe at 298 K. $^{1}H/^{1}H$ TOCSY [4], NOESY [5] and DQF COSY spectra were recorded on an unlabelled sample using a 5500 Hz spectral width in both dimensions. $^{1}H/^{15}N$ HSQC [6], $^{1}H/^{15}N$ 3D HMQC-NOESY, 3D HMQC-TOCSY [7], 3D HMQC-NOESY-HMQC [8] were acquired on a uniformly ^{15}N-labelled protein with spectral widths of 1500 Hz in nitrogen dimensions and 5500 Hz in the proton dimensions. A mixing time of 100 ms was used in the NOESY experiments and 55 ms in the TOCSY experiments. HNCO [9] and CBCACONH [10] spectra were acquired using a $^{13}C/^{15}N$ double-labelled sample. NMR spectra were processed using the program FELIX 2.10 (Biosym). $^{3}J_{HN\alpha}$ coupling constants were estimated from splittings in HMQC-J spectra [11]. Amide protons that were protected from exchange with solvent were identified by dissolving 10 mg of ^{15}N-labelled protein in D_2O buffer solution and recording an HSQC spectra.

3. Results and discussion

Assignments for the domain were obtained using the method of Wüthrich [12]. 3D HMQC-TOCSY spectra and 2D $^{1}H/^{1}H$ TOCSY and DQF-COSY experiments were used to identify spin systems which were then linked by through-space NOEs observed in 3D HMQC-NOESY and 3D HMQC-NOESY-HMQC spectra. Assignments were confirmed using CBCA-CONH spectra (Fig. 2) that link amide proton and nitrogen shifts to the CA and CB resonances of the preceding residue. Assignments for carbonyl resonances were obtained from HNCO spectra.

Assignments could not be obtained for residues K628 and K629. Subsequent analysis indicated that these residues are in a loop where signals are often weak.

Areas of secondary structure in the domain were identified from patterns of sequential NOEs, characteristic medium- and long-range NOEs, together with J coupling information. Two regions of the domain are linked by stretches of NH-NH NOEs, contain $\alpha N(i, i+3)$, $\alpha\beta(i, i+3)$ and $\alpha N(i, i+4)$ NOEs, and have small $^{3}J_{HN\alpha}$ coupling constants (Fig. 3). These properties are characteristic of a helical conformation. The presence of two helices is further supported by the protection of NH protons to exchange with D_2O within these regions, which is indicative of the formation of hydrogen bonds. The beginning and end of the helices can not be identified precisely at this stage and must await a full structure determination. Based on the information available, it appears that the helical regions run approximately from P579–N591 and from S631–L646. There are three regions of the protein, H595–L599, N609–I615 and I619–N626, that exhibit runs of strong αN NOEs and large $^{3}J_{HN\alpha}$ coupling constants indicative of extended strand (Fig. 3). A number of long-range backbone/backbone NOEs were observed in these regions that show that they form a 3 stranded antiparallel β sheet (Fig. 4). This is confirmed by the observed pattern of amide protection. Residues in the central strand

*Corresponding author. Fax: (44) (223) 336 445.

SSDI 0014-5793(95)00181-6

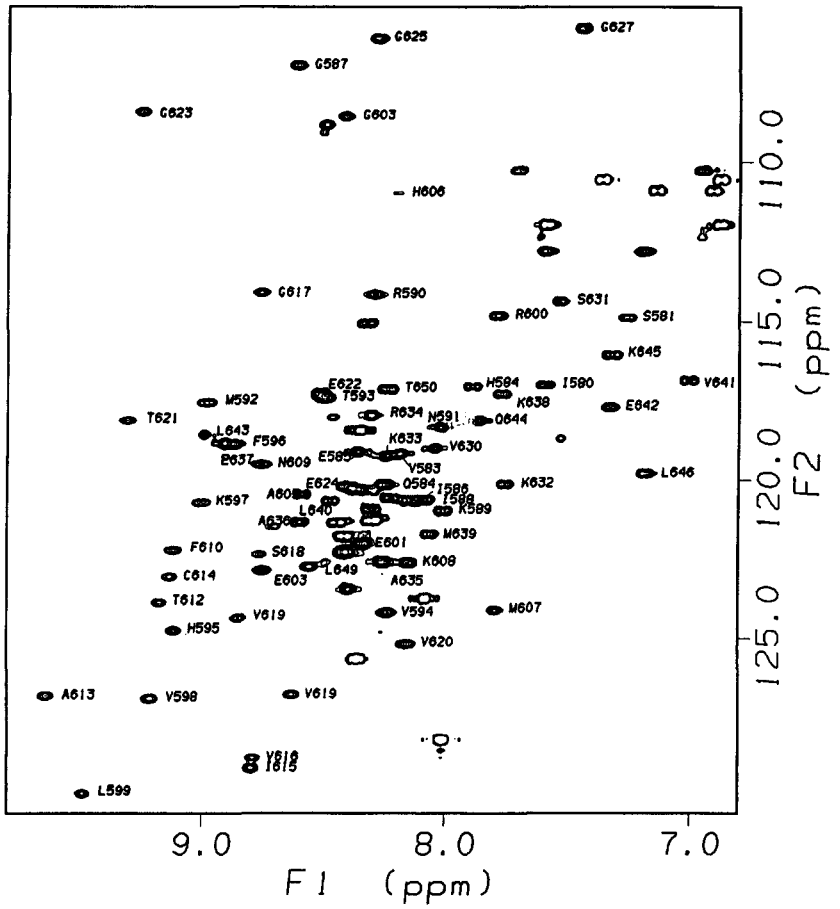

Fig. 1. $^{1}H/^{15}N$ HSQC spectrum of staufen double-stranded RNA binding domain. The assignments are indicated. The remaining peaks are from unstructured residues outside of the domain or from vector derived residues.

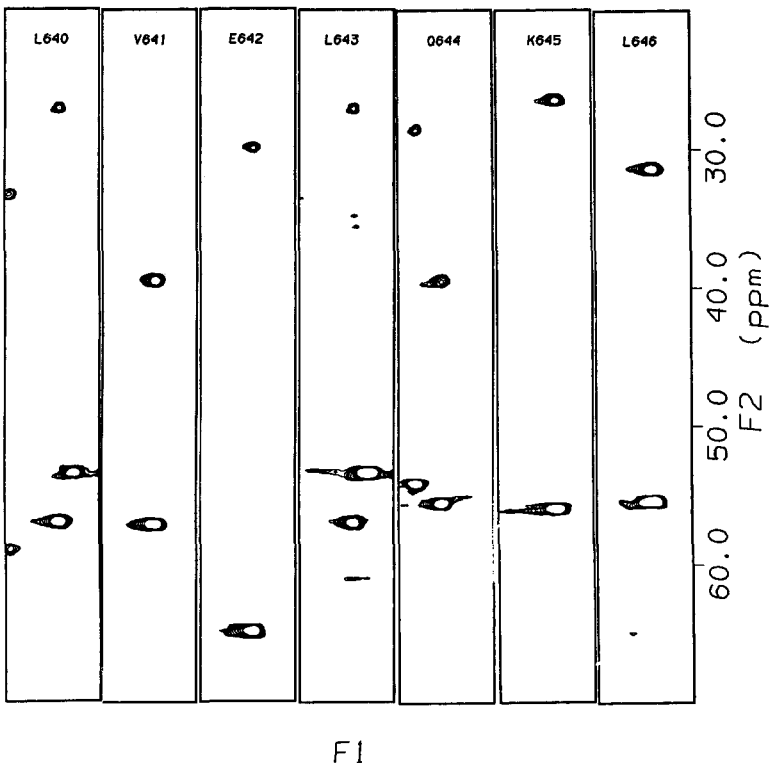

Fig. 2. Representative strips taken from the CBCACONH spectrum of the double-stranded RNA binding domain.

Table 1
Assignments for ^{1}H, ^{13}C, ^{15}N resonances in the third double-stranded RNA binding domain of staufen

	NH	^{15}N	CO	HA	CA	HB	CB	Others
P579			177.00	3.56	63.70	2.36	30.06	
I580	7.58	116.8	176.40	3.56	64.51	1.75	37.07	HG2 0.89, HG1 1.48, HD 0.72
S581	7.23	114.3	176.55	4.44	60.02	4.09	61.49	
Q582	8.22	120.1	177.15	4.17	58.28	2.24	30.50	HG 2.52, 2.33
V583	8.18	119.2	176.81	3.50	66.00	2.24	29.86	HG 1.13, 0.73
H584	7.89	116.5	117.62	4.36	55.43	3.36, 3.26	30.68	HD2 7.12, HE1 7.87
E585	8.37	119.3	178.45	4.31	58.30	2.27	27.83	
I586	8.12	120.6	177.25	3.67	64.22	1.81	36.07	HG2 0.93, HG1 1.07, HD 0.83
G587	8.60	107.0	175.23	3.74, 3.54	46.39			
I588	8.09	120.7	179.68	3.96	63.09	2.05	36.01	HG2 0.95, HG1 1.79, 1.35, HD 0.92
K589	8.01	121.3	177.61	4.12	57.86	2.06, 1.94	30.79	
R580	8.34	114.1	174.04	4.38	54.06	1.60	28.66	
N591	8.03	118.3	173.23	4.41	52.82	3.22, 2.75	35.46	
M592	8.99	117.7	174.17	4.74	53.01	1.94, 1.48	35.66	
T593	8.51	117.0	172.30	4.13	60.50	4.08	60.50	HG2 1.18
V594	8.19	124.1	174.52	4.80	59.70	1.93	32.29	HG 0.96, 0.21
H595	9.11	125.2	171.89	5.05	53.36	3.10, 2.94	32.65	HD2 6.96, HE1 8.01
F596	8.70	118.4	174.77	5.66	54.93	2.99	40.87	HD 7.24, HE 7.18, HZ 6.94
K597	8.98	120.4	173.88	4.78	53.36	1.80, 1.65	34.88	
V598	9.20	126.5	175.35	4.30	62.53	2.19	29.66	HG 1.09
L599	9.50	129.8	176.84	4.39	54.68	1.62	41.57	HG 1.70, HD 0.94, 0.92
R600	7.74	114.2	176.06	4.64	54.95	1.76, 1.48	31.23	
E601	8.39	122.2	174.10	4.69	52.89	1.79, 1.66	29.83	
E602	8.82	122.9	174.26	4.75	53.09	2.17, 1.94	31.43	
G603	8.39	108.75		4.63, 3.88				
P604			174.82	4.58	60.80	2.33, 2.01	30.64	
A605	8.63	120.2	177.70	3.89	53.92	1.47	16.66	
H606	8.05	111.0	174.42	4.68	54.82	3.44, 3.05	28.21	HD2 7.04, HE1 7.93
M607	7.68	125.2	172.22	4.61	53.04	1.91, 1.49	31.00	
K608	8.15	122.7	175.85	4.37	54.14	1.72, 1.23	32.05	
N609	8.80	119.5	172.70	5.07	50.83	2.65, 2.54	39.98	
F610	9.14	122.5	174.66	5.02	55.81	2.87	39.84	HD 7.09, HE 7.34, HZ 7.31
I611	8.86	124.2	175.17	5.16	58.83	1.86	36.80	HG2 0.92, HG1 1.40, 1.25, HD 0.78
T612	9.18	123.8	171.43	5.05	60.27	3.71	70.20	HG 1.22
A613	9.64	126.80	175.22	5.61	48.31	1.35	18.90	
C614	9.12	122.2		4.33	58.92	2.70		
I615	8.80	129.3	175.75	4.97	58.90	1.50	39.00	HG2 0.82, HG1 1.35, 1.05, HD 0.72
V616	8.78	128.4	175.75	4.91	59.20	2.23	32.18	HG 0.83, 0.78
G617	8.73	114.0	173.67	4.08, 3.75	45.80			
S618	8.78	122.4	172.75	4.48	56.80	4.05	62.54	
I619	8.63	126.9	173.27	4.04	56.05	2.00	38.11	
V620	8.16	125.2	176.11	5.33	59.02	1.90	33.24	0.85, 0.86
T621	9.31	118.2	172.14	4.85	58.34	4.31	70.88	HG2 1.07
E622	8.51	117.0	174.87	5.57	53.15	2.29, 2.03	31.85	
G623	9.26	108.0	176.19	4.67, 3.77	37.03			
E624	8.39	119.8	175.19	5.89	52.50	2.04, 2.27	31.45	
G625	8.26	105.9		4.36, 4.16	40.00			
N626			173.61	5.27	52.10	3.30, 2.97	35.98	
G627	7.42	105.6		4.12				
K628								
K629								
V630	8.06	119.0	176.21	3.89	64.09	2.04	29.60	HG 1.11
S631	7.52	114.2	174.95	3.88	58.65	4.08	58.92	
K632	7.75	119.8	177.65	3.83	59.29	1.72	30.47	
K633	8.22	118.9	175.41	3.77	59.32	2.03	30.38	
R634	8.36	118.9	178.86	4.20	57.72	2.14	28.09	
A635	8.24	122.4	179.97	3.85	54.23	1.58	16.27	
A636	8.58	120.9	177.74	4.24	53.38	1.64	16.60	
E637	8.90	119.0	179.57	3.92	58.70	2.47, 2.10	28.10	
K638	7.76	116.8	178.82	4.20	58.05	2.03	31.44	
M639	8.08	121.7	175.63	4.65	55.50	2.30	30.76	
L640	8.46	120.4	178.32	3.91	57.10	1.92, 1.62	39.52	HG 1.81, HD 0.00, 0.83
V641	7.00	117.1	178.17	3.70	64.93	2.36	30.15	HG 1.17, 0.99
E642	7.31	117.7	180.10	4.14	56.96	2.38, 2.27	27.22	
L643	9.01	118.6	179.33	4.11	55.58	2.07, 1.57	39.41	HG 1.90, HD 0.89, 0.84
Q644	7.87	118.1	175.58	4.23	56.00	2.26	26.58	
K645	7.30	115.9	176.46	4.30	55.46	2.06, 1.92	31.40	
L646	7.16	119.5		4.50	59.77	1.78, 1.45	46.16	HG 1.90, HD 0.94, 0.86

Spectra recorded at pH 6.8, 298 K.

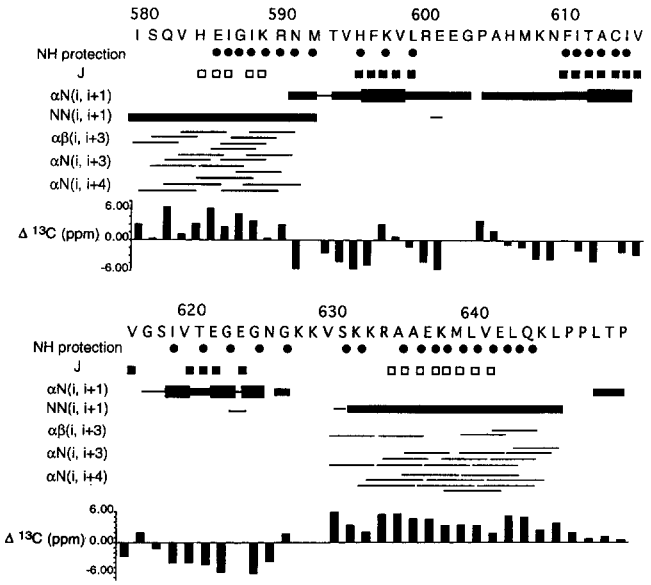

Fig. 3. Figure summarising the sequential, medium-range NOE, amide protection J coupling data and ^{13}C chemical shift data observed for the RNA binding domain. The strength of the NOE is indicated by the thickness of the line. Amide protons protected in a sample of the domain freshly dissolved in D_2O are indicated by filled circles. Residues with $^3J_{HN\alpha}$ couplings greater than 8 Hz are indicated by filled squares, and residues with $^3J_{HN\alpha}$ couplings smaller than 5 Hz are indicated by open squares. The variation of the Cα chemical shift from the random coil value minus the variation of the Cβ chemical shift from the random coil value is shown.

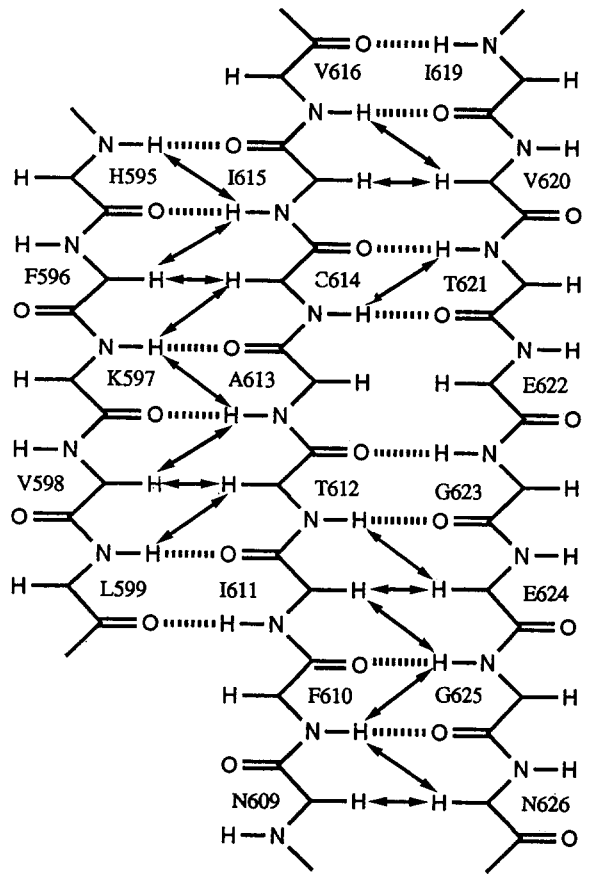

Fig. 4. Figure illustrating the inter-strand NOEs observed in the 3 stranded antiparallel β sheet. The absence of NOEs between parts of the second and third strands is the result of the similarity of the α proton chemical shifts of A613 and E622.

show continuous protection whilst residues in the peripheral strands show an alternating pattern of protection. The first two strands are linked by a 9 residue loop whereas the last two strands are linked by a hairpin turn. The CA and CB chemical shifts show the expected general trends in variation from random coil values that have been found for residues in the helix and sheet (Fig. 3) [13].

The dsRNA binding domain has an αββββα secondary structure arrangement. Structure information is available for two other RNA binding domains, the RNP motif [14] and the cold-shock domain that is homologous to eukaryotic Y box factors [15–17]. The RNP motif is also an α/β protein but has an βαββαβ secondary structure arrangement, whilst the cold-shock domain has a β barrel fold. Most of the residues that are conserved in the domain are located within regions of secondary structure. The most highly conserved sequence, GXGXXKXXXKXXAA (residues G623–A636 in this protein), span the end of the third strand of the β sheet and the start of the C-terminal helix. The conserved lysine residues in this region are unlikely to be important structurally and may well be involved in RNA binding. The nature of the RNA protein interaction will become clearer on completion of the determination of the three-dimensional structure of the domain.

Acknowledgements: M.B. and M.P. were supported by a ZENECA/MRC/DTI LINK program. D.St.J. is the recipient of a Wellcome Trust Senior Fellowship.

References

[1] Burd, C.G. and Dreyfuss, G. (1994) Science 265, 615–621.
[2] StJohnston, D., Brown, N.H., Gall, J.G. and Jantsch, M. (1992) Proc. Natl. Acad. Sci. USA 89, 10979–10983.
[3] Gibson, T.J., Thompson, J.D. and Heringa, J. (1993) FEBS Lett. 324, 361.
[4] Braunschweiler, L. and Ernst, R.R. (1983) J. Magn. Reson. 53, 521–528.
[5] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) J. Chem. Phys. 71, 4546–4553.
[6] Bodenhausen, G. and Ruben, D.L. (1980) Chem. Phys. Lett. 69, 185–188.
[7] Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) J. Magn. Reson. 86, 204–209.
[8] Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) J. Magn. Reson. 90, 420–425.
[9] Kay, L.E., Ikura, R., Tschudin, R. and Bax, A. (1990) J. Magn. Reson. 89, 496–514.
[10] Siomi, H., Matunis, M.J., Micael, W.M. and Dreyfuss, G. (1993) Nucl. Acids Res. 21, 1193–1198.
[11] Ramakrishnan, V. and White, S.W. (1992) Nature 358, 768–771.
[12] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[13] Spera, S. and Bax, A. (1992) J. Am. Chem. Soc. 113, 5490–5492.
[14] Nagi, K., Oubridge, C., Jenssen, T.H., Li, J. and Evans, P.R. (1990) Nature 348, 515.
[15] Schindelin, H., Marahiel, M.A. and Heinemann, U. (1993) Nature 364, 164–168.
[16] Schnuchel, A., Wiltscheck, R., Czisch, M., Herrier, M., Willimsky, G., Graumann, P., Marahiel, M.A. and Holak, T.A. (1993) Nature 364, 169–171.
[17] Schindelin, H., Jiang, W., Inouye, M. and Heineman, U. (1994) Proc. Natl. Acad. Sci. USA 91, 5119–5123.