Journal of Biomolecular NMR, 3 (1993) 535–546
ESCOM

J-Bio NMR 139

3D $^{13}C/^{1}H$ NMR-based assignments for side-chain resonances of *Lactobacillus casei* dihydrofolate reductase. Evidence for similarities between the solution and crystal structures of the enzyme

A. Soteriou[a], M.D. Carr[a], T.A. Frenkiel[b], J.E. McCormick[a], C.J. Bauer[b], D. Šali[a], B. Birdsall[a] and J. Feeney[a,]*

[a]*Laboratory of Molecular Structure and* [b]*MRC Biomedical NMR Centre, National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.*

Received 11 February 1993
Accepted 7 June 1993

Keywords: Dihydrofolate reductase, NMR; $^{13}C/^{15}N$ labelled protein

SUMMARY

^{13}C-based three-dimensional ^{1}H-^{1}H correlation experiments have been used to determine essentially complete ^{13}C and ^{1}H resonance assignments for the amino acid side chains of uniformly $^{13}C/^{15}N$ labelled *L. casei* dihydrofolate reductase in a complex with the drug methotrexate. Excellent agreement is observed between these assignments and an earlier set of partial assignments made on the basis of correlating nuclear Overhauser effect and crystal structure data, indicating that the tertiary structure of the enzyme is similar in solution and in the crystal state.

INTRODUCTION

Dihydrofolate reductase (DHFR) catalyses the reduction of dihydrofolate to tetrahydrofolate using NADPH as coenzyme. The enzyme is of considerable pharmacological interest since it is the target for a number of anticancer, antimalarial and antibacterial drugs, collectively known as antifolates (Blakley, 1985). Many unanswered questions still exist concerning the interactions between DHFR and its ligands; for example, little is known about the structural basis for the species selectivity of certain antifolate drugs, nor about the origins of the cooperativity between the substrate and coenzyme binding sites. In order to address questions of this type, we have been using NMR spectroscopy to examine selected binary and ternary complexes of *L. casei* DHFR, with the aim of determining their detailed three-dimensional (3D) structure in solution. The first step in this process is to obtain essentially complete sequence-specific ^{1}H, ^{15}N and ^{13}C resonance

*To whom correspondence should be addressed.

0925-2738/$ 10.00 © 1993 ESCOM Science Publishers B.V.

assignments, a task which is not straightforward since *L. casei* DHFR is a relatively large protein by NMR criteria (162 residues, M_r 18 300).

Recently, we published several studies that combined uniform ^{15}N labelling of *L. casei* DHFR with 3D $^{15}N/^{1}H$ TOCSY-HMQC (Marion et al., 1989a), NOESY-HMQC (Marion et al., 1989b) and HMQC-NOESY-HMQC (Frenkiel et al., 1990; Ikura et al., 1990) experiments (Carr et al., 1991). This approach enabled us to overcome the problems resulting from amide proton chemical-shift degeneracy and to make sequence-specific backbone resonance assignments for 146 of the 162 residues in the methotrexate (MTX) complex of DHFR. In addition, we were able to determine the secondary structure of the protein in solution and to show that it was similar to that observed in the crystal state (Bolin et al., 1982).

For proteins of the size of DHFR, the ability to trace out complete amino acid spin systems using ^{1}H NMR methods is severely limited, not only by signal overlap, but also by the absence of many possible cross peaks. This latter problem arises because the proton–proton coupling constants active in magnetisation transfer are small (3–15 Hz) compared to the linewidths of the ^{1}H resonances (> 20 Hz). Recently, it has been shown that these problems can be overcome for proteins of up to about 250 amino acids, by examining uniformly ^{13}C labelled proteins with 3D $^{13}C/^{1}H$ HCCH-COSY and HCCH-TOCSY experiments (Bax et al., 1990a,b; Ikura et al., 1991a), which make use of the large single-bond C-H (120–140 Hz) and C-C (35–45 Hz) couplings to obtain through-bond proton-to-proton correlations. The potential of this approach has been demonstrated in studies of the proteins interleukin 1β (153 residues) (Clore et al., 1990) and calmodulin (148 residues) (Ikura et al., 1991a,b), where these methods led to the determination of complete ^{13}C and ^{1}H resonance assignments for the side chains.

In this report we describe essentially complete ^{1}H and ^{13}C resonance assignments for the side chains of all the nonaromatic residues in the *L. casei* DHFR-MTX complex, based on analysis of HCCH-COSY and HCCH-TOCSY spectra from a uniformly $^{13}C/^{15}N$ labelled protein sample. The ^{1}H assignments of the aromatic residues have been reported previously (Birdsall et al., 1990). A comparison between the sequential assignments reported here and a previous set of partial assignments obtained by correlating NOE and crystal structure data (Hammond et al., 1986; Carr et al., 1991) allows us to comment on the tertiary structure of the protein in solution.

MATERIALS AND METHODS

The 99% ^{13}C enriched glucose and ^{15}N enriched ammonium sulphate were obtained from Cambridge Isotope Laboratories. 100% atom D_2O and methotrexate were purchased from Sigma. All other reagents were of A.R. quality.

Uniformly ^{13}C and ^{15}N labelled *L. casei* DHFR was prepared from an *E. coli* strain into which the gene for the *L. casei* enzyme had been cloned (NF1/pMT 702) (Andrews et al., 1985). The cells were grown on a minimal medium containing 3 g/l 99% ^{13}C enriched glucose, 1 g/l 99% ^{15}N enriched ammonium sulphate, 20 g/l potassium phosphate, 0.2 g/l magnesium sulphate, 50 mg/l ampicillin and 40 mg/l L-tryptophan (the latter was included in the growth medium since the *E. coli* strain used is auxotrophic for tryptophan). The purification of the protein was carried out using the same protocol as reported for unlabelled DHFR (Dann et al., 1976). The NMR experiments were carried out on 0.6-ml samples of 1.5 mM DHFR-MTX dissolved in a 100% D_2O, 500 mM potassium chloride and 50 mM potassium phosphate buffer at a pH* of 6.5 (pH*

values refer to the actual pH meter readings uncorrected for deuterium isotope effects). It should be noted that although the protein samples used in this study were uniformly labelled with both ^{13}C and ^{15}N only the former is required for the experiments described here.

Two 3D $^{13}C/^{1}H$ experiments were carried out to obtain the side-chain assignments described in this paper. Both were implemented on a Varian UNITY-600 spectrometer, using the pulse sequences described by Bax and co-workers (Bax et al., 1990a,b; Ikura et al., 1991a), with presaturation used to suppress the residual HDO signal. The NMR measurements were performed at a sample temperature of 35 °C. In the HCCH-TOCSY experiment (Bax et al., 1990b), carbon-carbon isotropic mixing was achieved by means of a DIPSI-3 sequence (Shaka et al., 1988) with a 25-ms mixing time and a radio-frequency field strength of 8.3 kHz. The HCCH-COSY experiment (Bax et al., 1990a) was of the constant-time variety (Ikura et al., 1991a), with a carbon evolution time of 7.8 ms. Both experiments were carried out in the phase-sensitive mode (States et al., 1982; Marion and Wüthrich, 1983) with the carrier positioned at the centre of the region of interest in all three dimensions. Sixteen transients were averaged for each increment with 512 points per transient.

The HCCH-TOCSY spectrum was recorded using the TPPI method of quadrature detection (Marion and Wüthrich, 1983), acquiring 128 × 128 increments with spectral widths of 8.33 ppm in the proton dimensions and 71.6 ppm in the carbon dimension, sufficient to accommodate all relevant signals without folding. This spectrum took approximately 65 h to record. In the constant-time HCCH-COSY experiment, the same proton spectral widths were used but the carbon spectral width was reduced by a factor of three. This reduced spectral width results in folding in the carbon dimension, but allows the resolution to be improved. In this case, the quadrature detection method of States et al. (1982) was used and 192 × 50 increments were acquired over about 40 h.

In the folded spectrum each ^{13}C (F_2) slice corresponds to three possible ^{13}C chemical shifts: the actual measured frequency and this value plus or minus the ^{13}C spectral width. However, because of the availability of the nonfolded spectrum it was relatively straightforward to determine which of these possibilities was the correct one for each set of ^{1}H-to^{1}H correlations.

The 3D datasets were transformed, displayed and plotted using a SUN SPARC-330 workstation utilising software written in-house. In order to improve the resolution in the spectra, the number of data points in F_1 and F_2 was initially doubled using linear prediction. The time-domain matrices were then zero-filled to 512 × 256 × 1024 points, resulting in final spectra consisting of 256 × 128 × 512 real points. In addition, mild resolution enhancement was achieved by applying a $\pi/2.5$ shifted sine-squared apodisation function in all dimensions.

RESULTS

In 3D $^{13}C/^{1}H$ HCCH-COSY and HCCH-TOCSY spectra (Bax et al., 1990a,b; Ikura et al., 1991a), through-bond ^{1}H-to-^{1}H correlations are characterised by the ^{1}H (F_1) and ^{13}C (F_2) chemical shifts of the proton-carbon pair from which the magnetisation originates and by the ^{1}H (F_3) shift of the proton to which magnetisation is transferred. Hence, the spectra can be thought of as consisting of a series of 2D ^{1}H COSY (Marion and Wüthrich, 1983; Aue et al., 1986) or TOCSY (Braunschweiler and Ernst, 1983; Davies and Bax, 1985) slices (F_1/F_3), separated according to ^{13}C (F_2) chemical shifts. The identification of amino acid spin systems and their assignment to specific

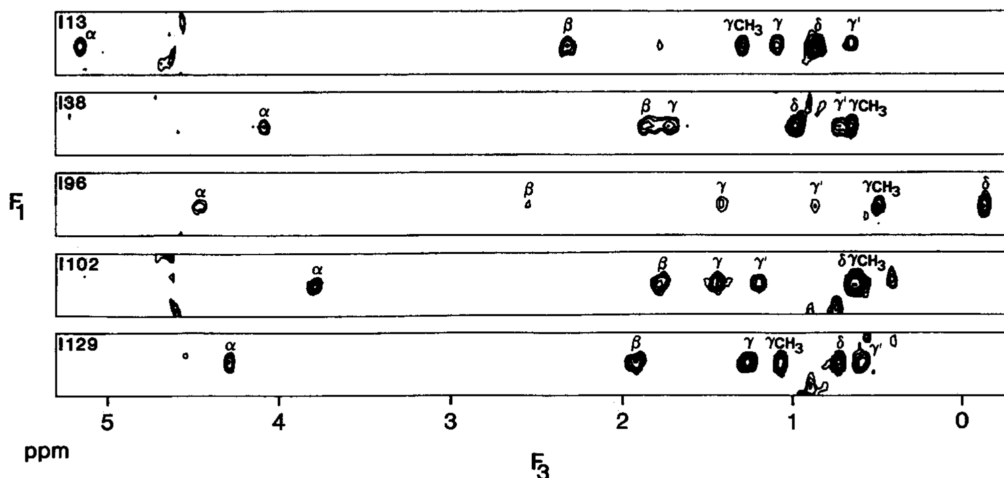

Fig. 1. A series of F_1/F_3 strips from the HCCH-TOCSY spectrum of the DHFR-MTX complex taken at the δ-carbon shifts (F_2) of I13 (15.11 ppm), I38 (16.23 ppm), I96 (11.20 ppm), I102 (10.64 ppm) and I129 (16.79 ppm). The labelled cross peaks correspond to through-bond correlations spanning the full length of the isoleucine side chains, from the δ-methyl protons (F_1) to γ-methyl, γ-, γ'-, β- and α-protons (F_3).

types or classes of residues, is therefore based upon the observation of the same correlation patterns expected in 2D COSY and TOCSY spectra (Wüthrich, 1986), with HCCH-COSY experiments principally being used to distinguish direct from relayed connectivities. In the 3D spectra, though, a cross peak corresponding to an αCH-to-βCH correlation, for example, will typically appear in a different ^{13}C (F_2) slice from that of the corresponding βCH-to-αCH cross peak, because the α- and β-carbons usually have different ^{13}C chemical shifts. Consequently, in order to trace out complete amino acid spin-system patterns in HCCH-COSY or HCCH-TOCSY spectra it is necessary to move between ^{13}C (F_2) planes. For instance, after identifying an αCH-to-βCH cross peak in the F_2 slice corresponding to the α-carbon shift, it would be necessary to determine the corresponding β-carbon shift in order to find correlations originating from the βCH. A convenient way to solve this problem is to make use of spectra sliced along the starting proton shift (F_1). For example, in the F_2/F_3 slice corresponding to the ^1H shift (F_1) of the βCH one can identify the expected correlation to the αCH and thereby read off the corresponding β-carbon shift.

In the case of DHFR-MTX, using the HCCH-COSY and HCCH-TOCSY spectra it proved possible to obtain side-chain assignments for the majority of the residues. This process is illustrated by a description of the assignment of isoleucine, leucine and lysine resonances, which are generally considered to be the most difficult signals to identify in all types of spectra.

Isoleucine assignments

Identification of resonances from the isoleucine residues in HCCH-TOCSY and HCCH-COSY spectra of the DHFR-MTX complex was fairly straightforward. For example, the spectrum showed cross peaks corresponding to correlations between the δCH_3 protons and αCH, βCH, γCH, γ'CH and γCH_3 protons for all five isoleucines in F_2 (^{13}C) slices of the HCCH-TOCSY spectrum corresponding to the δ-carbon shifts, as shown in Fig. 1. In addition, the direct δCH_3-to-γCH and γ'CH connectivities were observed in the HCCH-COSY spectrum. Thus, it was a

relatively simple process to determine the α, β, γ and γCH_3 carbon shifts for each isoleucine, using the strategy outlined previously and then to identify further ^{1}H-to-^{1}H correlations. It should be noted that the characterisation of each isoleucine spin system is highly overdetermined, since we observed cross peaks corresponding to between 31 and 41 of the 44 possible direct and relayed through-bond connectivities (14 HCCH-COSY and 30 HCCH-TOCSY connectivities). This situation contrasts sharply with that found in 2D COSY and TOCSY and 3D ^{15}N/^{1}H TOCSY-HMQC spectra of DHFR-MTX (Carr et al., 1991), where due to extensive signal overlap and the absence of many expected cross peaks, side-chain assignments could be obtained for only two of the isoleucines.

The α-proton resonances of the five isoleucines in the protein are well resolved from one another (Table 1) so there is no problem in matching the side-chain assignments obtained in the present work to the sequence-specific backbone assignments determined previously (Carr et al., 1991).

Leucine assignments

Although DHFR contains a relatively large number of leucine residues (13), it was fairly easy to fully characterise all of them by examining the HCCH-COSY and HCCH-TOCSY spectra. The signals from the leucine spin systems are particularly clear in the α-carbon planes (F_2) of the HCCH-TOCSY spectrum, where for 11 of the residues it was possible to identify all the expected correlations starting at the α-proton (some of these are shown in Fig. 2); for the remaining two leucines all the correlations except one could be detected.

Previously, the δCH_3, δ'CH_3 and γCH leucine signals had been identified from a 2D COSY spectrum of a DHFR-MTX sample in which the valine methyl groups were fully deuterated (Birdsall et al., 1984). However, for all leucine residues except Leu^{131}, it proved impossible to observe through-bond correlations to the corresponding amide, α- or β-protons in COSY, TOCSY or ^{15}N/^{1}H TOCSY-HMQC spectra, because of signal overlap and missing cross peaks (Carr et al., 1991). Consequently, most of the sequence-specific ^{1}H assignments for leucine side chains reported previously (Hammond et al., 1986; Carr et al., 1991) were based on correlating NOEs observed in NOESY spectra of DHFR-MTX with those predicted from consideration of the crystal structure of the related DHFR-MTX-NADPH complex (Bolin et al., 1982).

The complete leucine side-chain assignments obtained from HCCH-COSY and HCCH-TOCSY spectra have now been compared with previous, solely NMR-based, sequential assignments for α- and β-protons (Hammond et al., 1986; Carr et al., 1991; Soteriou et al., unpublished results). This comparison has allowed us to confirm the previous sequence- specific assignments made for all the leucine residues in DHFR-MTX.

Recently we have examined the 2D COSY spectrum of a DHFR-MTX complex formed with selectively deuterated enzyme incorporating stereoselectively deuterated leucine, i.e. (2*S*,4*R*) [5,5,5-$^2H_3$]-leucine (Ostler et al., 1993). The stereospecific assignments of the leucine methyl resonances were made by noting which cross peaks were absent in this spectrum compared to that obtained using nondeuterated enzyme. The spectrum of the selectively deuterated sample showed cross peaks involving leucine methyls only for the 4-*pro-S* methyl groups.

Clearly, the ^{1}H stereospecific assignments obtained in this way can be used to make stereospecific assignments for the leucine Cδ carbons and these are given in Table 1.

TABLE 1
^{13}C AND ^{1}H RESONANCE ASSIGNMENT FOR SIDE CHAINS IN THE METHOTREXATE COMPLEX WITH *L. casei* DHFR

Residue	C_α	H_α	H_α'	C_β	H_β	H_β'	C_γ	H_γ	H_γ'	C_δ	H_δ	H_δ'	C_{other}	H_{other}
1 Thr	64.34[a]	4.19[a]		70.36[a]	4.43[a]		23.49[a]	1.22[a]						
2 Ala	51.99	5.99[a]		23.99	1.53[a]									
3 Phe	55.74	5.99		42.24	3.53	3.26								
4 Leu	55.36	6.18		49.24	2.05	1.33	28.49	2.34		27.36	1.16	0.53	32.61 (Cδ')	
5 Trp	59.87	5.56		34.74	3.55	3.55								
6 Ala	51.49	5.44		23.49	1.40									
7 Gln	54.37	6.17		37.29	2.33	1.93	33.99	2.39	2.13					
8 Asp	53.99	4.69		42.99	4.21	2.85								
9 Arg														
10 Asp	56.49	5.04		44.24	2.97	2.68								
11 Gly	46.99	4.58	3.51											
12 Leu	58.49	4.37		45.74	1.93	1.70	28.61	1.70		25.74	0.93	0.80	27.86 (Cδ')	
13 Ile	62.61	5.21		43.11	2.36		27.36	1.13	0.70	14.99	0.89		19.61 (Cγ')	1.33 (γ')
14 Gly	46.99	4.35	3.81											
15 Lys	56.99	4.39		36.86	1.71	1.45	26.74	1.27	1.14	30.74	1.71	1.57	43.61 (Cε)	2.94(ε) 2.94(ε')
16 Asp	57.12	4.28		41.99	2.93	2.43								
17 Gly	47.19	3.97	3.47											
18 His	55.37	4.98		32.24	3.25	3.25								
19 Leu	54.74	4.23		43.61	1.53	0.39	29.11	1.41		26.86	0.66	0.46	25.36 (Cδ')	
20 Pro	63.74	4.35		32.49	2.45	1.81	29.65	1.35	1.33		3.13	3.13		
21 Trp	56.80	4.69		31.74	2.51	2.33								
22 His	57.44	4.91		33.49	3.17	2.92								
23 Leu	52.49	4.81		46.24	1.83	1.05	28.49	1.05		25.12	0.65	0.07	26.36 (Cδ')	
24 Pro	67.62	3.77		33.49	2.33	1.93								
25 Asp	59.87	4.29		44.74	3.03	2.28								
26 Asp	58.12	5.37		45.74	2.55	2.31								
27 Leu	59.74	3.90		41.86	1.81	1.08	27.99	1.48		26.74	0.63	0.02	22.74 (Cδ')	
28 His	58.74	4.67		29.72	3.51	3.37								
29 Tyr	63.24	4.17		40.24	3.39	3.09								
30 Phe	63.44	3.59		40.49	3.51	2.63								
31 Arg														
32 Ala	57.05	3.96		20.74	1.37									
33 Gln														
34 Thr	64.34	4.19		72.74	3.72		23.24	0.27						
35 Val	66.62	3.81		33.49	2.06		22.99	1.00	1.00				22.99 (Cγ')	
36 Gly	46.99	4.21	3.91											
37 Lys														
38 Ile	63.74	4.13		39.61	1.87		29.86	1.77	0.77	16.24	1.03		20.74 (Cγ')	0.70 (γ')
39 Met	55.38	5.25			2.39		32.87	2.27	2.14					
40 Val	62.74	5.09		34.24	1.80		22.37	0.71	0.63				23.49 (Cγ')	
41 Val	59.87	5.57		39.37	2.35		26.24	1.25	0.97				20.24 (Cγ')	
42 Gly	45.87	4.69	4.21											
43 Arg	61.49	3.57		32.49	1.50	1.27	27.99	1.05	1.05	44.76	2.85	2.49		
44 Arg	60.94	3.95		30.74	1.95	1.85	29.12	1.73	1.61	44.74	3.17	3.17		
45 Thr	68.87	3.75		70.37	4.04		22.49	1.11						
46 Tyr	62.74	2.89		42.12	2.79	2.79								

TABLE 1 (continued)

Residue	C_α	H_α	H_α'	C_β	H_β	H_β'	C_γ	H_γ	H_γ'	C_δ	H_δ	H_δ'	C_{other}	H_{other}
47 Glu	59.87	3.59		30.74	2.09	1.85	39.12	2.81	2.31					
48 Ser	60.62	4.38		65.49	4.05	3.89								
49 Phe	54.12	4.47		38.68	2.81	2.81								
50 Pro	66.24	4.41		33.62	2.43	2.00	28.99	2.07	2.07	52.68	4.13	3.97		
51 Lys	56.49	4.28		36.49	1.65	1.55	25.99	1.30	1.30	30.74	1.65	1.65	43.61 (Cε)	2.97(ε) 2.97(ε')
52 Arg	53.74	4.39		33.49	1.37	1.33	27.37	1.05	1.05	44.74	2.69	2.69		
53 Pro														
54 Leu	54.74	4.35		42.49	1.27	0.93	27.99	1.15		24.11	−0.07	0.36	27.99 (Cδ')	
55 Pro	64.37	4.53		35.12	2.39	2.09	29.12	2.09	1.97	52.62	3.77	3.47		
56 Glu	59.54	3.89		28.49	2.37	2.31	39.12	2.19	2.13					
57 Arg	55.99	4.73		33.57	1.73	1.39	29.66	1.25	1.19	44.76	3.21	2.93		
58 Thr	64.99	4.41		70.49	4.15		22.99	1.16						
59 Asn	64.24	5.06		41.24	3.13	2.29					6.89	8.00		
60 Val	62.12	4.82		34.18	1.87		23.49	0.61	0.37				23.24 (Cγ')	
61 Val	61.49	4.69		34.74	1.56		21.74	0.37	−0.12				21.24 (Cγ')	
62 Leu	54.49	4.76		44.74	1.55	0.73	29.24	1.17		27.36	0.43	0.66	25.49 (Cγ')	
63 Thr	60.37	4.75		70.99	3.73		19.49	0.93						
64 His	58.19	4.74			3.43	3.32								
65 Gln	58.12	4.13		29.74	2.05	1.97	34.74	2.37	2.37					
66 Glu	61.12	3.31		31.24	1.89	1.85	37.99	2.21	1.99					
67 Asp	54.37	4.65		41.93	2.83	2.63								
68 Tyr	62.05	4.16		41.37	2.92	2.61								
69 Gln	56.26	4.31		32.49	1.89	1.73	35.24	2.27	2.27					
70 Ala	52.74	4.35		21.80	1.02									
71 Gln	59.30	4.09		33.57	2.05	2.05	35.24	2.39	2.39					
72 Gly	46.99	4.25	3.71											
73 Ala	51.49	4.84		22.37	1.13									
74 Val	63.74	3.98		34.12	1.71		21.74	0.61	0.45				21.74 (Cγ')	
75 Val	63.68	4.65		34.12	1.97		23.99	0.85	0.79				23.99 (Cγ')	
76 Val	59.87	4.43		36.99	2.19		23.49	0.67	0.53				20.74 (Cγ')	
77 His	55.99	5.49		31.87	3.53	3.05								
78 Asp	54.59	4.65		45.24	2.93	2.93								
79 Val	67.74	3.43		33.49	1.83		23.49	0.61	0.61				23.49 (Cγ')	
80 Ala	57.12	4.33		19.12	1.53									
81 Ala	56.43	4.35		20.62	1.66									
82 Val	68.87	3.53		32.24	2.49		25.74	1.11	0.79				23.99 (Cγ')	
83 Phe	63.74	4.54		39.74	3.35	3.04								
84 Ala	56.99	4.25		19.62	1.59									
85 Tyr	64.24	4.14		39.74	3.28	3.20								
86 Ala	57.62	4.00		19.49	1.75									
87 Lys	60.36	4.07		34.11	2.02	1.97	26.86	1.68	1.58	30.74	1.73	1.73	43.61 (Cε)	3.01(ε) 3.01(ε')
88 Gln	57.62	4.09		30.74	1.87	1.87	35.74	2.45	2.33					
89 His	55.49	4.81		30.87	3.50	2.74								
90 Pro	66.24	4.37		33.49	2.42	2.07	29.37	2.42	2.27	53.05	4.17	3.53		
91 Asp	55.93	4.57		41.99	2.75	2.75								
92 Gln	56.62	4.88		35.14	2.27	2.09	37.49	2.37	2.27					
93 Glu	57.54	4.42[a]		34.69[a]	2.33[a]	2.33[a]	32.99	2.77	2.69					
94 Leu	56.49	4.99[a]		46.49[a]	1.66[a]	1.66[a]	29.11	1.48		26.99	0.78	0.78	26.99 (Cδ')	

TABLE 1 (continued)

Residue	C_α	H_α	H_α'	C_β	H_β	H_β'	C_γ	H_γ	H_γ'	C_δ	H_δ	H_δ'	C_{other}	H_{other}
95 Val	62.62	4.84		36.37	1.88		23.49	0.95	0.88				23.62 (Cγ')	
96 Ile	60.99	4.52		37.99	2.58		29.61	1.46	0.89	11.14[a]	−0.09[a]		21.86 (Cγ')	0.54 (γ')
97 Ala	53.18	5.98		21.74	1.73									
98 Gly	44.24	4.21	2.25											
99 Gly	46.99	4.05	3.75											
100 Ala	58.74	4.02		20.12	1.59									
101 Gln	60.37	4.15		29.12	2.11	2.11	35.74	2.55	2.41					
102 Ile	62.54	3.83		36.36	1.81		27.86	1.50	1.23	10.61	0.66		19.86 (Cγ')	0.69 (γ')
103 Phe	61.55	4.05		39.12	2.73	2.54								
104 Thr	69.43	3.75		70.49	4.24		23.49	1.27						
105 Ala	56.49	4.11		19.09	1.33									
106 Phe	60.99	4.99		44.24	3.91	2.99								
107 Lys	62.11	4.09		33.49	2.02	2.02	25.11	1.65	1.40	31.24	1.80	1.75	43.61 (Cε)	3.09(ε) 3.09(ε')
108 Asp	57.62	4.81		42.62	2.83	2.56								
109 Asp	56.49	4.95		45.99	3.12	2.69								
110 Val	63.24	3.89		34.24	1.28		20.12	0.45	0.39				26.24 (Cγ')	
111 Asp	56.49	4.99		45.37	2.99	2.75								
112 Thr	63.74	5.57		74.93	3.93		22.99	1.41						
113 Leu	55.36	5.03		43.11	1.90	1.43	28.24	0.91		24.61	−0.95	0.30	25.74 (Cδ')	
114 Leu	55.86	5.35		44.24	2.60	1.63	28.74	1.77		27.36	0.96	1.07	27.99 (Cδ')	
115 Val	60.37	4.32		35.99	0.61		21.24	0.61	−0.01				23.49 (Cγ')	
116 Thr	62.12	4.92		70.44	4.04		23.62	0.56						
117 Arg	55.99	4.73		33.49	1.73	1.37	29.99	1.19	1.27	44.74	3.21	2.91		
118 Leu	58.11	4.53		43.86	1.88	1.07	32.99	0.93		24.36	−0.33	−0.51	27.24 (Cδ')	
119 Ala	55.37	4.06		20.80	1.37									
120 Gly	45.87	4.21	3.55											
121 Ser	58.18	5.06		65.99	3.58	3.58								
122 Phe	58.74	4.76		43.84	3.35	2.89								
123 Glu	57.62	4.59		32.49	2.05	2.05	37.99	2.33	2.23					
124 Gly	47.37	4.12	4.03											
125 Asp	55.49	4.95		45.30	3.05	2.65								
126 Thr	64.37	4.68		73.49	3.76		22.37	1.23						
127 Lys	56.49	5.07		36.86	1.91	1.69	27.36	1.53	1.28	30.74	1.70	1.70	43.61 (Cε)	2.99(ε) 2.99(ε')
128 Met	56.34	4.11		30.59	2.01	1.41	33.01	1.53	1.41					
129 Ile	61.74	4.34		37.99	1.96		26.74	1.30	0.63	16.74	0.77		21.49 (Cγ')	1.11 (γ')
130 Pro	64.98	4.49		33.62	2.29	1.85	29.67	2.12	2.02	51.99	3.85	3.61		
131 Leu	54.36	4.35		46.49	0.50	0.21	27.36	1.25		24.99	−0.02	0.44	28.49 (Cδ')	
132 Asn	53.18	5.01		37.49	3.05	2.72								
133 Trp	63.12	3.79		30.24	3.21	2.97								
134 Asp	58.24	4.78		41.99	2.86	2.75								
135 Asp	57.24	4.76		42.49	2.77	2.54								
136 Phe	58.74	5.21		46.49	3.35	3.28								
137 Thr	62.12	4.89		73.24	3.98		22.93	1.21						
138 Lys	58.74	3.61		31.24	1.13	−0.74	25.74	0.57	0.09	30.49	1.29	1.21	42.49 (Cε)	2.63(ε) 2.53(ε')
139 Val	64.37	4.23		34.62	2.13		22.99	0.95	0.91				21.74 (Cγ')	
140 Ser	59.12	4.62		66.99	3.83	3.69								
141 Ser	59.34	5.31		66.55	3.61	3.46								
142 Arg	57.54	4.79		35.24	1.93	1.93	29.12	1.65	1.51	45.24	3.33	3.17		

TABLE 1 (continued)

Residue	C_α	H_α	H_α'	C_β	H_β	H_β'	C_γ	H_γ	H_γ'	C_δ	H_δ	H_δ'	C_{other}	H_{other}
143 Thr	64.30	4.78		72.12	3.80		22.93	0.89						
144 Val	63.24	3.99		35.24	0.67		22.37	0.93	0.85				22.37 (Cγ')	
145 Glu	57.54	4.49		32.34	1.84	1.84	38.24	2.21	2.08					
146 Asp	57.24	4.76		42.49	2.77	2.54								
147 Thr	66.74	3.92		70.49	4.13		23.31	1.29						
148 Asn	51.04	5.13		40.04	3.34	2.84								
149 Pro	64.94	4.63		35.81	2.39	2.07	26.30	1.97	1.75	51.49	3.65	3.60		
150 Ala	55.37	3.96		19.49	1.38									
151 Leu	55.36	4.36		42.99	2.00	1.74	28.99	1.60		27.36	1.14	0.84	25.11 (Cδ')	
152 Thr	66.49	4.30		70.87	4.21		23.99	1.11						
153 His	54.24	5.74		32.37	2.69	2.12								
154 Thr	62.12	5.01		73.24	3.69		24.12	1.06						
155 Tyr	59.12	4.97		40.64	2.79	2.61								
156 Glu	55.99	5.33		35.49	2.39	2.01	37.49	2.27	2.13					
157 Val	63.74	5.25		35.74	2.03		21.87	0.95	0.95				21.87 (Cγ')	
158 Trp	56.62	5.93		32.99	3.45	3.13								
159 Gln	55.99	5.62		34.74	2.23	2.04	35.74	2.57	2.45					
160 Lys	59.24	3.77		34.61	1.81	1.57	26.49	1.36	0.83	30.74	1.49	1.49	43.61 (Cε)	2.83(ε) 2.83(ε')
161 Lys	59.24	4.14		34.74	1.83	1.61	27.36	1.44	1.23	31.24	1.69	1.69	43.61 (Cε)	2.97(ε) 2.97(ε')
162 Ala	55.37	4.15		21.93	1.33									

The ^{13}C and ^1H chemical shifts (ppm) are referenced to TSP (3-trimethylsilylpropionate) and DSS (2,2-dimethyl-2-silapentane-5-sulfonate), respectively. The leucine CH_3 signals have been stereospecifically assigned (Ostler et al., 1993): H_δ = 4-*Pro-R*; H_δ' = 4-*Pro-S*.

[a] These signals appear as double signals: the origin of this nonequivalence is under investigation.

Lysine assignments

The ends of lysine side chains tend to be found on the surface of proteins because of the presence of the positively charged, terminal NH_3^+ group. One consequence of this is that the ε and δ ^1H resonances tend to have chemical shifts close to the random-coil values. This situation also appears to apply for the ^{13}C signals from the ε carbons, since in a single F_2 (^{13}C) slice of the HCCH-TOCSY spectrum of DHFR-MTX we observed well-resolved εCH_2-to-αCH correlations for seven of the nine lysines, as well as a group of overlapping cross peaks corresponding to εCH_2-to-δCH, -δ'CH, -γCH, -γ'CH, -βCH and -β'CH connectivities, as shown in Fig. 3. Analogous correlations for an eighth lysine residue were seen in a neighbouring F_2 slice of the HCCH-TOCSY spectrum. After identifying these correlation patterns in the ε carbon slices of the HCCH-TOCSY and HCCH-COSY spectra, it proved fairly straightforward to trace out the signals for the complete spin systems of Lys^{15}, Lys^{51}, Lys^{87}, Lys^{107}, Lys^{127}, Lys^{138}, Lys^{160} and Lys^{161}. For example, in the HCCH-TOCSY spectrum we observe cross peaks corresponding to all possible correlations beginning at the α-protons of these residues.

The αCH resonances of Lys^{15}, Lys^{51}, Lys^{127}, Lys^{138}, Lys^{160} and Lys^{161} are well resolved from one another (Table 1) and consequently, sequence-specific assignments for these residues could be made simply by matching the α-proton shifts from the HCCH-COSY and HCCH-TOCSY spectra with those obtained by sequential assignment of the protein backbone (Carr

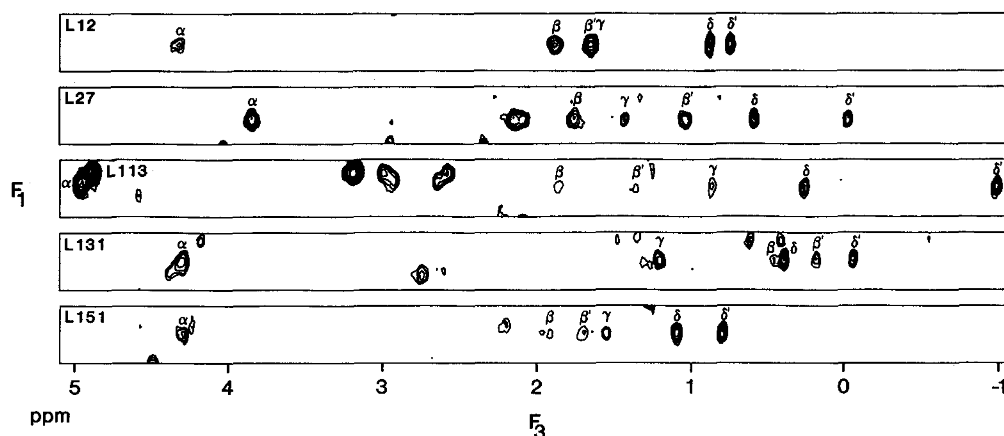

Fig. 2. Representative F_1/F_3 strips from the HCCH-TOCSY experiment taken at the α-carbon shifts (F_2) of Leu[12] (58.73 ppm), Leu[27] (59.85 ppm), Leu[113] (55.37 ppm), Leu[131] (54.26 ppm) and Leu[151] (55.37 ppm). The labelled cross peaks are typical of the correlations seen from the α-protons (F_1) of all 13 leucine residues in the DHFR-MTX complex, with magnetisation being transferred along the full length of the side chains.

et al., 1991). However, this method fails for Lys[87] and Lys[107] since both the α- and β-protons of these residues have overlapping proton chemical shifts. Thus, the sequence-specific assignments given in Table 1 for these two residues initially relied on NH-to-γCH and -γ'CH NOEs, observed in the 2D NOESY (Jeener et al., 1979; Macura et al., 1981) and 3D $^{15}N/^{1}H$ NOESY-HMQC (Marion et al., 1989b) spectra. However, these assignments have recently been confirmed by triple-resonance HNCA experiments (Soteriou et al., unpublished results). In the case of Lys[37], no assignments have been reported here, since it was not possible to identify a ninth amino acid spin system with the correlation pattern expected for a lysine residue, probably because both the proton and carbon chemical shifts for the γ- and δ-positions have identical, or very similar, values.

DISCUSSION

Analysis of 3D $^{13}C/^{1}H$ HCCH-COSY and HCCH-TOCSY spectra recorded from a uniformly $^{13}C/^{15}N$ labelled DHFR-MTX sample has enabled us to obtain complete ^{13}C and ^{1}H resonance assignments for the side chains of essentially all the residues in the protein. In earlier work it had not been possible to obtain complete side-chain assignments for long-chain amino acids such as Ile, Leu and Lys using DQF-COSY, TOCSY and $^{15}N/^{1}H$ TOCSY-HMQC spectra, principally because many of the expected cross peaks were absent for reasons described previously. Clearly, in addition to the extra resolution gained from the third frequency dimension, the great strength of HCCH-type experiments is the substantial increase in sensitivity, which arises from using the large, single-bond C-H (120–140 Hz) and C-C (35–45 Hz) couplings to establish proton-to-proton correlations, rather than relying on the small values for two- or three-bond proton-proton coupling constants (3–15 Hz).

Previously, sequence-specific assignments for the γ-, δ- and δ'-protons of the 13 leucine residues in DHFR-MTX were obtained by correlating NOE and crystal structure data (Hammond et al.,

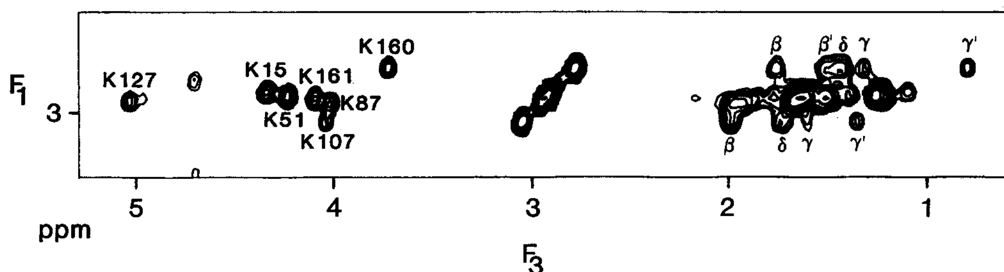

Fig. 3. An F_1/F_3 slice from the HCCH-TOCSY spectrum of the DHFR-MTX complex corresponding to the ε-carbon shift (F_2 = 43.64 ppm) of seven of the nine lysines in the protein. The cross peaks, identified by residue number, correspond to intense εCH (F_1) to αCH (F_3) correlations, whilst the cluster of peaks between 1 and 2 ppm (F_3) arises from ε-to-δ, -γ and -β correlations, as indicated for Lys^{107} and Lys^{160}. The identification of the ε-to-α connectivities provided an excellent starting point for tracing out the complete lysine spin systems.

1986; Carr et al., 1991). In addition, this approach was used to make stereospecific assignments for the methyl resonances from 12 of the 13 leucine residues (confirmed later by examining selectively deuterated DHFR incorporating (2*S*,4*R*)[5,5,5-$^2H_3$]-leucine (Ostler et al., 1993)). We have now been able to confirm all these previous leucine assignments using solely protein-sequence and NMR data. Thus, since the original leucine assignments, including the stereospecific assignments for the methyls, were dependent upon networks of long-range interresidue NOEs, this agreement indicates that the overall tertiary structure of *L. casei* DHFR must be similar in solution and in the crystal state.

In earlier NMR studies (Carr et al., 1991) it was shown that the secondary-structure elements found in the crystal structure (four α-helices and an eight-stranded β-sheet (Bolin et al., 1982)) remain essentially intact in solution. The comparison of the leucine residue assignments obtained from the crystal-based method with those from the sequential assignment method, discussed above, indicates that the tertiary structure is also retained. In addition, interresidue NOEs involving many other assigned side-chain protons (particularly aromatic resonances (Birdsall et al., 1990)) also confirm the observation that the crystal and solution structures are globally similar. However, it should be noted that small differences in structure would not be detected by this approach because only a qualitative comparison of observed and calculated NOEs is used for the crystal-structure based assignments.

The only crystal-structure data available for an *L. casei* DHFR complex are those for the ternary DHFR-MTX-NADPH complex (Bolin et al., 1982). However, it has proved possible to use these data to make crystal-based assignments for several other binary and ternary complexes, which have subsequently been examined using sequential-assignment methods, and in all cases there is good agreement between crystal-based assignments and the sequential assignments obtained without use of crystallographic data. It is encouraging that the crystal structure for one complex can be so useful for assigning the spectrum of a different ligand complex formed with the same enzyme. This allows a series of different complexes to be examined by NMR with relative ease, while such an examination is often difficult using X-ray crystallography.

Such crystal-based approaches to assignments could be important for larger proteins for which conventional sequential-assignment methods are difficult to apply.

ACKNOWLEDGEMENTS

The NMR measurements were made using the facilities at the MRC Biomedical NMR Centre, NIMR, Mill Hill. D.S. acknowledges the award of a Lloyds Tercentenary Fellowship.

REFERENCES

Andrews, J., Clore, G.M., Davies, R.W., Gronenborn, A.M., Gronenborn, B., Kalderon, D., Papadopoulos, P.C., Schäfer, S., Sims, P.F.G. and Stancombe, R. (1985) *Gene*, **35**, 217–222.

Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976) *J. Chem. Phys.*, **64**, 2229–2246.

Bax, A., Clore, G.M., Driscoll, P.C., Gronenborn, A.M., Ikura, M. and Kay, L.E. (1990a) *J. Magn. Reson.*, **87**, 620–627.

Bax, A., Clore, G.M. and Gronenborn, A.M. (1990b) *J. Magn. Reson.*, **88**, 425–431.

Birdsall, B., Feeney, J., Griffiths, D.V., Hammond, S., Kimber, B., King, R.W., Roberts, G.C.K. and Searle, M. (1984) *FEBS Lett.*, **175**, 364–368.

Birdsall, B., Arnold, J.R.P., Jimenez-Barbero, J., Frenkiel, T.A., Bauer, C.J., Tendler, S.J.B., Carr, M.D., Thomas, J.A., Roberts, G.C.K. and Feeney, J. (1990) *Eur. J. Biochem.*, **191**, 659–668.

Blakley, R.L. (1985) In *Folates and Pterins*, Vol. 1 (Eds, Blakley, R.L. and Benkovic, S.J.) Wiley, New York, pp. 191–253.

Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. and Kraut, J. (1982) *J. Biol. Chem.*, **257**, 13650–13662.

Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.*, **53**, 521–528.

Carr, M.D., Birdsall, B., Frenkiel, T.A., Bauer, C.J., Jimenez-Barbero, J., Polshakov, V.I., McCormick, J.E., Roberts, G.C.K. and Feeney, J. (1991) *Biochemistry*, **30**, 6330–6341.

Clore, G.M., Bax, A., Driscoll, P.C., Wingfield, P.T. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 8172–8184.

Dann, J.G., Ostler, G., Bjur, R.A., King, R.W., Scudder, P., Turner, P.C., Roberts, G.C.K., Burgen, A.S.V. and Harding, N.G.L. (1976) *Biochem. J.*, **157**, 559–571.

Davies, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.*, **107**, 2820–2821.

Frenkiel, T.A., Bauer, C.J., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.*, **90**, 420–425.

Hammond, S.J., Birdsall, B., Searle, M.S., Roberts, G.C.K. and Feeney, J. (1986) *J. Mol. Biol.*, **188**, 81–97.

Ikura, M., Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 9020–9022.

Ikura, M., Kay, L.E. and Bax, A. (1991a) *J. Biomol. NMR*, **1**, 299–304.

Ikura, M., Kay, L.E., Krinks, M. and Bax, A. (1991b) *Biochemistry*, **30**, 5498–5504.

Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.

Macura, S., Huong, Y., Suter, D. and Ernst, R.R. (1981) *J. Magn. Reson.*, **43**, 259–281.

Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989b) *J. Am. Chem. Soc.*, **111**, 1515–1517.

Ostler, G., Soteriou, A., Moody, C.M., Khan, J.A., Birdsall, B., Carr, M.D., Young, D.W. and Feeney, J. (1993) *FEBS Lett.*, **318**, 177–180.

Shaka, A.J., Lee, C.J. and Pines, A. (1988) *J. Magn. Reson.*, **77**, 274–293.

States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) *J. Magn. Reson.*, **48**, 286–292.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.