Journal of Peptide Science, Vol. 2, 91–105 (1996)

The Solution Conformational Features of Two Highly Homologous Antigenic Peptides of Foot-and-mouth Disease Virus Serotype A, Variants A and USA, Correlate with their Serological Properties

MONICA PEGNA[1,2], HENRIETTE MOLINARI[2,3], LUCIA ZETTA[2], G. MELACINI[2,*] WILLIAM A. GIBBONS[4], FRED BROWN[5], DAVE ROWLANDS[6], E. CHAN[1] and PAOLO MASCAGNI[1]

[1] Department of Peptide Chemistry, Italfarmaco Research Centre, Milan, Italy
[2] NMR laboratory, Istituto di Chimica delle Macromolecole, CNR, Milan, Italy
[3] Istituto Policattedra, University of Verona, Italy
[4] Department of Pharmaceutical Chemistry, School of Pharmacy, University of London, UK
[5] US Department of Agriculture, Agricultural Research Service, Plum Island Animal Disease Center, New York, USA
[6] Department of Molecular Sciences, Wellcome Foundation, UK

Received 1 May 1995
Accepted 29 September 1995

The solution structure of a peptide corresponding to the VP1 region 141–160 of foot-and-mouth disease virus (FMDV) serotype A variant USA has been studied by NMR and computer calculations and compared with the results from a study on a highly homologous peptide deriving from serotype A, variant A. The two peptides differ in their serological behaviour and contain the immunodominant epitope of the virus which partly overlaps with its receptor binding region. Distance constraints, derived both from 2D and 3D homonuclear NMR and 2D-heteronuclear NMR experiments, were combined with DG calculations to yield 50 structures. After refinement through EM and restrained molecular dynamics simulations the selected structures shared several general features. In particular the 151–158 region was a helix in all cases while a large loop similar to that found in peptide A but comprising less residues and stabilized by an H-bond between the side chains of D147 and S150 was found in the majority of structures. A further loop, common to all structures, was identified around the RGD sequence (145–147). This was different from that found in the corresponding region of peptide A as were the conformations of the individual residues within the RGDX sequence.

The different structural features shown by the two peptides were rationalized in terms of the S148 (peptide A) to F148 (peptide USA) mutation. The second mutation, that at position 153 (L in A, P in USA) did not appear to affect the structure of the peptide significantly although the different dimensions of the loop in the central region and the type of H-bond stabilizing it could be potentially ascribed to this second mutation.

All criteria used pointed to different structural features for the two peptides consistent with their serological behaviour.

Keywords: FMDV; NMR spectroscopy; RGD (Arg-Gly-Asp); NOE constraints; structure–activity correlation

* Present Address: Department of Chemistry, University of California, S. Diego, La Jolla, Ca 92093-0343, USA
Address for correspondence: P. Mascagni, Italfarmaco Research Centre, Via Lavoratori 54, Cinisello Balsamo, 20092 Milan, Italy; Tel. 39-2-6443-3090, Fax 39-2-66011579.

CCC 1075-2617/96/020091-15

INTRODUCTION

In the accompanying paper the solution structure of a peptide (peptide A) corresponding to the 141–160 sequence of FMDV, serotype A, was calculated using NMR parameters coupled to computer calculations. In this paper a similar study is described for peptide USA, a sequence corresponding to the 141–160 region of a different FMDV variant, the USA variant, which, despite only two mutations at positions 148 and 153 (Figure 1), has been shown to have a serological behaviour different from that of peptide A [1].

To obtain representative models of the solution structure of this peptide, NOE-derived interproton distances were used. It should be emphasized, however, that the NMR parameters were suggestive of conformational averaging. Thus the NOE-derived interproton distances were used to draw *qualitative* conclusions about the conformational space adopted by peptide USA and, above all, to combine, at the atomic level, the conformational differences between these two peptides which had been inferred from preliminary CD spectroscopy studies [2, 3] and serological studies [1].

MATERIALS AND METHODS

Materials

The peptide was synthesized as a carboxyamide derivative on a solid support (MBHA resin) using t-BOC chemistry. After HF cleavage the peptide was purified by RP-HPLC (purity > 97%) and identified by FAB-MS (expected 2039; found M + H 2040).

CF_3–CD_2–OH (TFE-d2) was obtained by distillation of a 1 : 1 (v/v) mixture of CF_3–CD_2–OD (Cambridge Isotope Laboratories Ltd, Cambridge, UK) and H_2O. The fractions used for the NMR experiments had boiling points ranging from 73 to 75 °C.

NMR Spectroscopy

NMR spectra were obtained on either a Bruker AM-500 spectrometer or a Bruker AMX 600 MHz, using

141-142-143-144-145-146-147-**148**-149-150-151-152-**153**-154-155-156-157-158-159-160

Gly-Ser-Gly-Val-Arg-Gly-Asp-**Phe**-Gly-Ser-Leu-Ala-**Pro**-Arg-Val-Ala-Arg-Gln-Leu-Pro (USA)

Gly-Ser-Gly-Val-Arg-Gly-Asp-**Ser**-Gly-Ser-Leu-Ala-**Leu**-Arg-Val-Ala-Arg-Gln-Leu-Pro (A)

Gly-Ser-Gly-Val-Arg-Gly-Asp-**Leu**-Gly-Ser-Leu-Ala-**Pro**-Arg-Val-Ala-Arg-Gln-Leu-Pro (B)

Gly-Ser-Gly-Val-Arg-Gly-Asp-**Ser**-Gly-Ser-Leu-Ala-**Ser**-Arg-Val-Ala-Arg-Gln-Leu-Pro (C)

Figure 1 Amino acid sequence of the 141–160 region of four antigenic variants (A, B, C and USA) of FMDV, serotype A, sub-type 12, strain 119. The four peptides differ for the residues at positions 153 and 148. In the text the residues were numbered from 1 to 20.

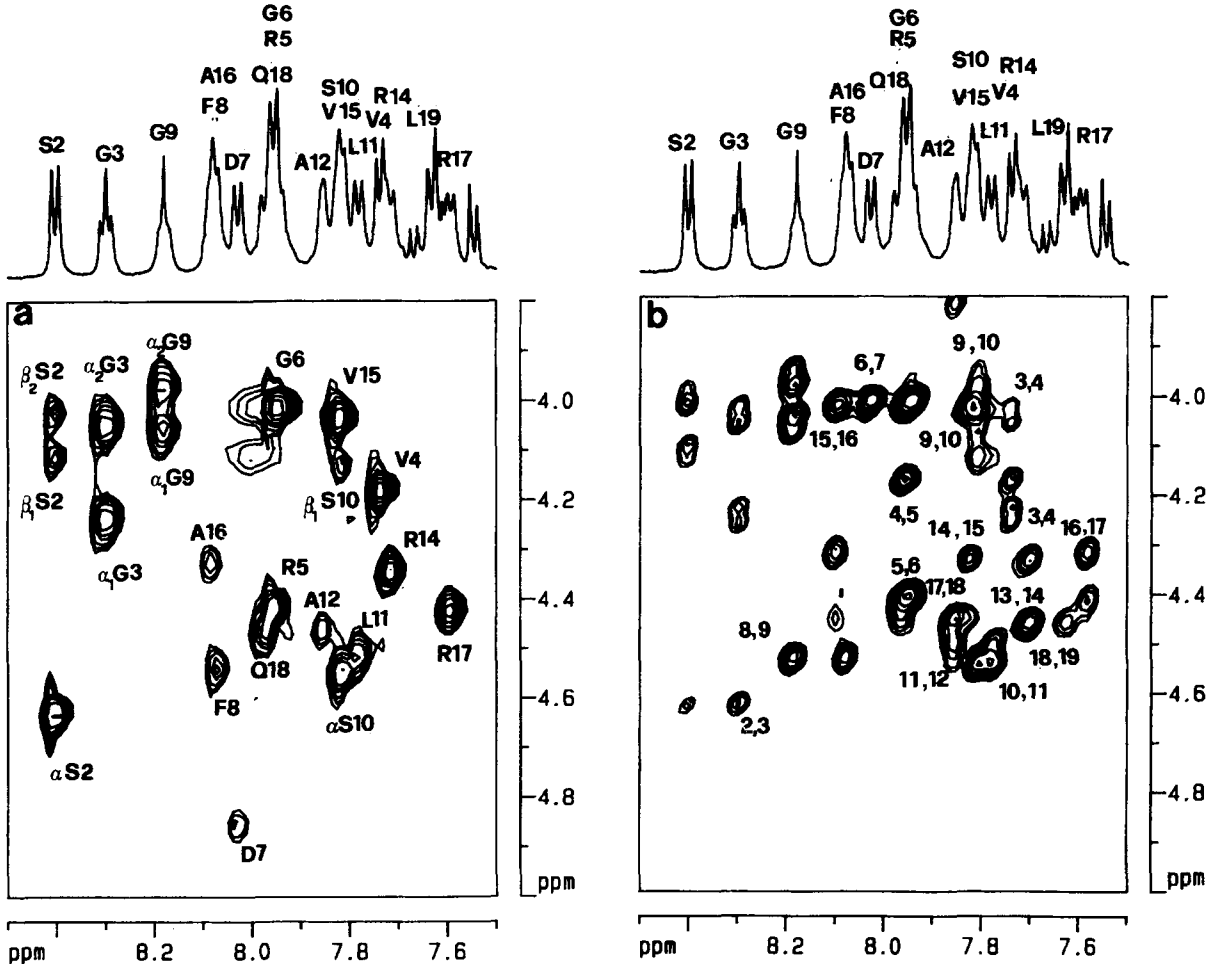


Figure 2 ^{1}H-NMR spectra of peptide USA in TFE-d2 solutions (2 mM). NH-αCH region of: clean-TOCSY (a) and NOESY (b). The spectra were acquired on a Bruker AM-500 spectrometer, the spectral width was 6250 Hz and 512 t_1 increments were implemented over 2048 points in t_2 dimension. Data matrix was zero filled to 1024 in F_1 and a $\pi/2$ shifted sine-bell window was applied in both dimensions. The mixing times used for the two sequences were 50 and 450 ms, respectively.

2 mM samples of peptide in TFE-d2 unless otherwise specified.

2D homonuclear experiments were performed as described for peptide A (accompanying paper). The 3D NOE-HOHAHA [4] spectrum was recorded on a 10 mM sample of peptide in TFE-d_2. FIDs (512 words, 16 scans preceded by 2 dummy scans) were acquired with the carrier positioned in the middle of the spectrum. The solvent peak was suppressed by irradiation during the relaxation delay of 1.5 s. The entire acquisition procedure was repeated for 256 t_2 times and 128 t_1 times, thus resulting in a total measuring time of eight days. The spectral width in all domains was 6250 Hz. The mixing time to allow for NOE magnetization transfer was 300 ms. A MLEV-17 sequence of 50 ms with two trim pulses at the beginning and the end of the mixing time was used for the homonuclear Hartman–Hahn transfer. No zero-filling was performed and a cosine-bell window was used in all dimensions.

The heteronuclear 2D ^{1}H-^{13}C SQC-NOESY [5] experiment was performed at 305 °K on a 10 mM sample of peptide using the pulse sequence described by Molinari *et al.* [6]. The delays were based on a $J_{HC\alpha}$ of 143 Hz; the spin-lock purging pulses were set to 3.5 ms and τ_m to 450 ms. GARP1 was employed as a decoupling scheme, using a ^{13}C 90° pulse of 44.5 μs. The acquisition times were 0.141 s in t_2 and 0.026 s in t_1; 512 transients were collected for each t_1 increment. The spectrum was processed using a cosine-bell window function.

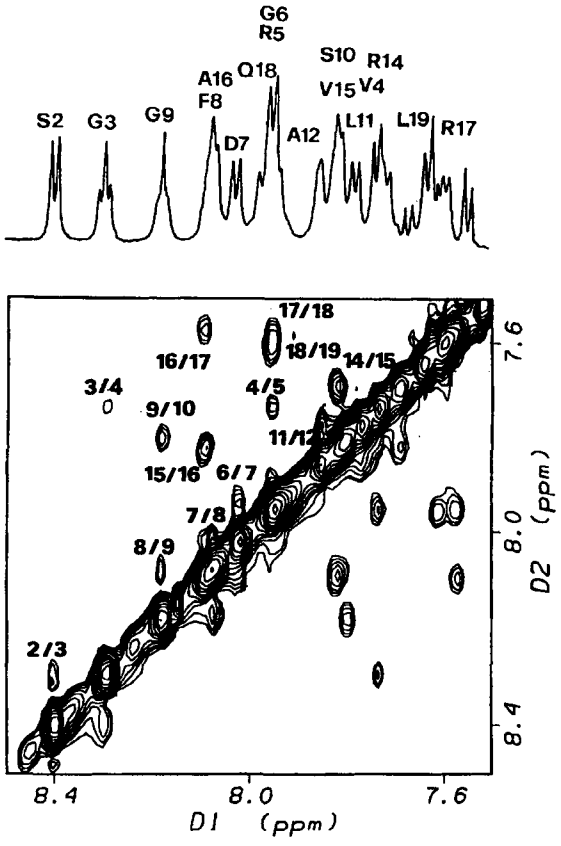


Figure 3 Amide region of the 2D-NOESY spectrum of peptide USA where NH–NH NOEs are shown. The spectrum was acquired on a Bruker AMX-600 spectrometer, the spectral width was 7500 Hz and 512 t_1 increments were implemented over 2048 points in t_2 dimension. Data matrix was zero filled to 1024 in F_1 and a $\pi/2$ shifted sine-bell window was applied in both dimensions. A mixing time of 200 ms was employed.

Molecular Modelling

All calculations were carried out using essentially the same approach as described in the accompanying paper.

RESULTS

2D and 3D-NMR Assignments of Peptide USA

The majority of sequence-specific and sequential assignments were made at 500 and 600 MHz using a combination of TOCSY, NOESY and ^1H–^{13}C SQC-NOESY experiments (Figures 2 and 3; Tables 1 and 2) [7, 8]. Spectral overlap (e.g. the NHs of R5 and G6 and those of S10 and L11) did not allow, however, the completion of the assignment of the proton spectrum.

To complete the latter a 3D homonuclear NOE-HOHAHA experiment [4] was carried out at 500 MHz. This permitted the separation of the resonances overlapping in conventional 2D spectra as shown in Figure 4 for the amide protons of R5 and G6 (Table 1). Furthermore, the 3D experiment provided an additional independent method for the crude evaluation of some interproton distances according to the procedure discussed below.

$J_{N\alpha}$-coupling constants were measured both at 500 and 600 MHz using a DQFCOSY experiment (Table 3). τ_c was measured from a heteronuclear ^1H–^{13}C SQC-NOESY experiment [5, 6, 9] using as reference the distance of 0.25 nm separating the β and γ protons of proline. A $\tau_c = 0.9$ ns was thus obtained.

Internuclear Distance Evaluation from NMR Data

The influence of spin diffusion on the NOE intensities is negligible when $\omega\tau_c$ is close to the NOE crossover value ($\omega\tau_c = 1.2$) [10, 11]. Furthermore when NOE data collected in the mixing time range 80–600 ms were used together with a correlation time of 0.9 ns to measure interproton distances for Gramicidin S, agreement was found with the X-ray structure of the cyclic peptide [12]. In the case of the peptide studied here τ_c was also 0.9 ns and NOE data were collected using mixing times ranging from 200 to 450 ms. This was used to conclude that the observed NOE intensities could be used to derive interproton distances (Figure 5). In the case of the homonuclear NOESYs at 500 and 600 MHz, distances were obtained using the first rate approximation method [7] (Figure 5).

An approximate measure of the α5N6, α9N10, $\alpha7\beta10$ and α17N18 distances was obtained from the 3D spectrum using the procedure described by Vuister *et al.* [4] (Figure 5).

NMR Temperature Studies

1D spectra were recorded in the temperature range 290–320 K and at 5 K intervals. TOCSY experiments were carried out at the two extreme temperatures. Of the 17 temperature coefficients thus measured (Table 3), 6 had values less than 4 p.p.b./K (G6, S10, L11, R14, R17 and Q18) and were assigned to hydrogen bonded NHs, whereas 3 residues (F8, G9, A16) had their NH temperature coefficients larger than approximately 6 p.p.b./K which was probably consistent with their being solvent-exposed. The remaining values (S2, G3, V4, R5, D7, A12, V15

Table 1 ^{1}H Chemical Shifts (p.p.m.) of Peptide USA in TFE at 305 K

Residue	NH	αH	βH	γH	Other
G1					
S2	8.40	4.62	4.23/4.09		
G3	8.29	4.22/4.02			
V4	7.73	4.17	2.25	2.09	
R5	7.94	4.38	2.00	1.85/1.75	$^{\delta}CH_2 = 3.29$ $^{\epsilon}NH = 7.10$
G6	7.94	4.00			
D7	8.02	4.90	2.98		
F8	8.07	4.51	3.25		2,6H = 7.35 3,5H = 7.36 4H = 7.42
G9	8.18	4.04/3.95			
S10	7.80	4.53	4.10/3.96		
L11	7.77	4.50	1.78	1.77	$^{\delta}CH_3 = 1.02/0.95$
A12	7.87	4.43	1.49		
P13	–	4.44	2.39/1.96	2.19/2.07	$^{\delta}CH_2 = 3.80/3.72$
R14	7.69	4.31	2.05	1.85/1.75	$^{\delta}CH_2 = 3.24$ $^{\epsilon}NH = 7.10$
V15	7.82	4.00	2.25	1.06/1.03	
A16	8.08	4.30	1.52		
R17	7.56	4.40	2.05/1.95	1.76/1.84	$^{\delta}CH_2 = 3.29$ $^{\epsilon}NH = 7.10$
Q18	7.97	4.35	1.80/1.52	2.50	
L19	7.61	4.80	1.83/1.76	1.75	$^{\delta}CH_3 = 1.02$
P20	–	4.47	2.32/2.09	2.85	$^{\delta}\&CH_2 = 3.87/3.69$

Table 2 ^{13}C Chemical Shifts (p.p.m.) of Peptide USA in TFE at 305 K

Residue	$^{\alpha}C$	$^{\beta}C$	$^{\gamma}C$	$^{\delta}C$	Other
G1	41.0				
S2	[a]	61.6			
G3	42.7				
V4	61.2	29.8	17.8		
R5	54.6	27.8	24.7	40.5	
G6	42.9				
D7	[a]	34.8			
F8	57.5	36.5			2,6Ar: 128.9 4Ar: 127.1 3,5Ar: 128.6
G9	43.2				
S10	–	61.6			
L11	60.5	40.1	24.7	19.7/20.1	
A12	51.5	13.8			
P13	62.3	28.7	24.9	47.5	
R14	55.2	27.8	24.7	40.5	
V15	61.8	29.8	17.8		
A16	51.1	15.5			
R17	54.6	27.8	24.7	40.5	
Q18	53.4	27.0	31.4		
L19	[a]	40.1	24.7	21.4/21.8	
P20	62.3	28.9	24.2	47.4	

[a] Correlations not defined since the corresponding proton was under the HDO peak.

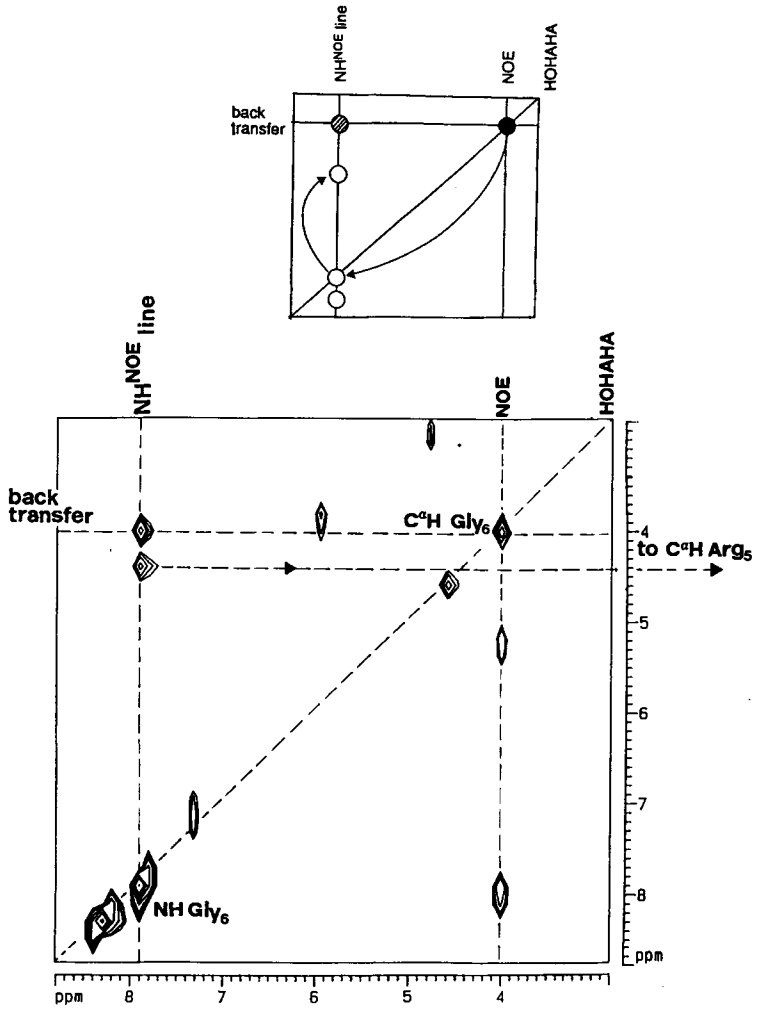


Figure 4 3D NOE-HOHAHA spectrum of a 10 mM solution of peptide USA in TFE-d_2 at 305 K; NH–αCH region. The spectrum was acquired on a Bruker AM-500 spectrometer; the spectral width was 6250 Hz, 128 t_1 and 256 t_2 increments were implemented over 2048 points in t_3 dimension.

The top panel shows a schematic representation of a 3D $\omega 3$ cross-section, where the diagonal is the intersection with the HOHAHA plane while the horizontal and vertical lines carry the back-transfer and NOE cross-peaks respectively. The bottom panel shows the αCH5–NH5 and αCH5-NH6 cross-peaks in the $\omega 3$ cross-section, taken at αCH6 frequency. These cross-peaks completely overlapped in the 2D-NMR spectra.

and L19) ranged from 4.1 to 5 p.p.b./K and might reflect less defined hydrogen bonds.

Patterns of NMR Data Indicating the Existence of Regular Secondary Structure Elements

The NOE data observed in the RGD region (Figures 2, 3 and 5) showed the following correlations: αH5–NH6 (strong), NH6–NH7 (strong) and αH5–NH7 (weak). Although these NOEs were consistent with a β-turn of type II around residues 4 to 7, the existence of a turn involving residues 5–8 could not be excluded since the crucial αH6–NH8 could have been obscured by spectral overlapping. However a temperature coefficient of 4.1 p.p.b./K found for NH7, which contrasted with that of 6.6 p.p.b./K for residue 8, suggested that the β-turn involving residues 4 and 7 was a more likely explanation for the NOE data. Both

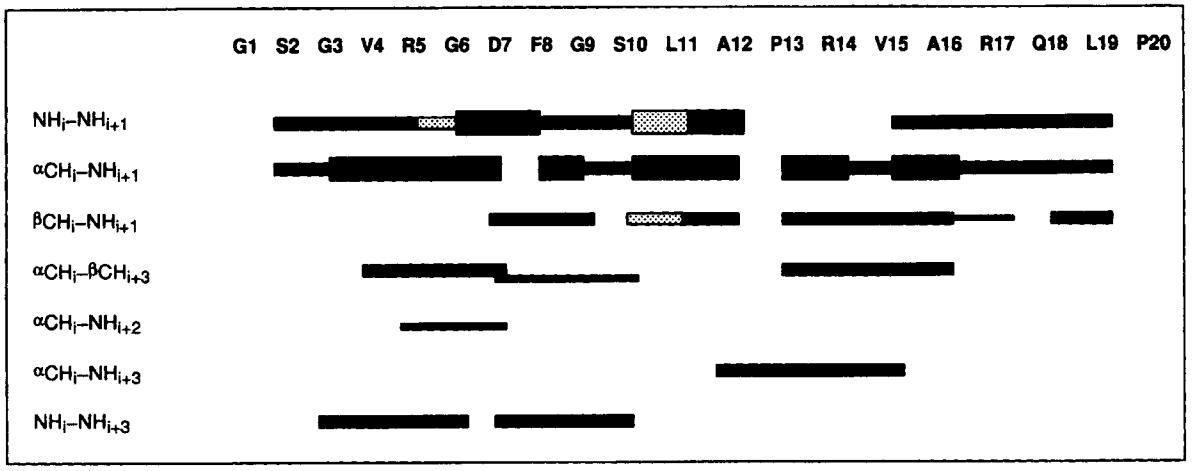


Figure 5 Summary of main chain NOE connectivities observed at 500 and 600 MHz. Gray areas of sequential NOEs indicate the presence of more than one overlapping peak. Strong, medium and weak NOEs are represented by line of different thickness. Some of the long range backbone–side chain connectivities are omitted for clarity reasons.

Table 3 Vicinal Coupling Constants and Temperature Coefficients of Peptide USA

Residue	$J_{NH-\alpha}$/Hz[a]	$\Delta\delta_{NH}/\Delta T$/p.p.b.
G1		
S2	7.4	4.8
G3	7.2/7.2[b]	4.4
V4	6.6	4.9
R5	6.6	4.9
G6	7.2/7.2[b]	3.8
D7	7.4	4.1
F8	7.1[b]	6.6
G9	7.1/7.1[b]	5.9
S10	7.3	3.0
L11	7.3	3.3
A12	3.6	5.2
P13	-	
R14	7.0	1.9
V15	6.1	5.0
A16	7.2[b]	7.4
R17	7.0	2.3
Q18	7.1[b]	2.4
L19	7.8	4.1
P20	-	-

[a] Vicinal coupling constants at 298 K.
[b] Measured from COSYDQF (600 MHz).

prediction studies for the RGD sequence and experimental findings for RGD-containing peptides [13] were consistent with this conclusion.

The temperature coefficient of the amide proton of G6 was 3.8 p.p.b./K and assigned the latter to a H-bonded proton. This, together with the NOE discussed above, could be interpreted as indicating the existence of two consecutive γ-turns, 4–6 and 5–7. The latter was recently proposed for a model FMDV peptide based on MD calculations [14].

The following data assigned the 7–12 region to a helix: (1) NH_i–NH_{i+1} NOEs (Figure 5); (2) the αH7–βH10 correlation; (3) all β–N NOEs expected for helical structures; and (4) the small NH temperature coefficients found for S10 and L11. It should be noted, however, that several other NOE constraints typical of compact α-helix or 3_{10} helix were not observed.

The NH12–δ13 NOE assigned a *trans* conformation to the alanine–proline amide bond while the NH14–δ13 NOE and the NH–R14 temperature coefficient of 1.9 p.p.b./K suggested that the latter was involved in a turn, either β or γ. The β-turn seemed to be a less likely explanation because, in this assumption, proline would occupy the less favoured position $i+2$ of the turn [16]. Furthermore the α11–NH14 and NH12–NH14 NOEs typical of β-turns were not detected. It was therefore concluded that the NH of R14 was hydrogen bonded to CO–A12, thus defining a γ-turn.

The NOE pattern assigned the 14–20 region to a helix (Figure 5). The temperature coefficients of the NH protons of residues 17, 18 and 19 (Table 3) were consistent with either a 3_{10} helix beginning with the CO of R14 or an α-helix starting with the CO of P13.

Simulations

Ninety-three NOE constraints (60 inter-residue and 23 inter-residue constraints) were used to generate structures by distance geometry calculations [16, 17]. Most of the 50 structures thus obtained had an

Table 4 E_{pot} (kcal) and $< r_{viol} >$ (Å) of the 23 Structures Selected after DG Calculations

Structure	E_{pot} (before RMD)	E_{pot} (after RMD)	$< r_{viol} >$
1	182	172	0.09
2	187	127	0.11
3	169	154	0.11
4	139	82	0.09
5	132	125	0.10
6	127	113	0.09
7	158	121	0.09
8	154	138	0.12
9	185	139	0.09
10	188	119	0.11
11	141	124	0.13
12	170	135	0.09
13	183	118	0.10
14	181	113	0.09
15	176	123	0.10
16	145	100	0.10
17	144	103	0.10
18	190	130	0.09
19	176	120	0.10
20	171	118	0.10
21	153	105	0.12
22	127	100	0.11
23	185	115	0.11

inverse γ-turn around P13, small restraint energies and high potential energies. The latter decreased after the subsequent EM [18] calculations. The structures were classified according to their (1) total potential energy, E_{tot}, (2) distance restraint energy, E_{res}, (3) number of distance restraint violations, N_{viol} (a violation was counted when a given distance exceeded the experimental input by 0.4 Å), (4) average distance violation $< r_{viol} >$ [19], and (5) consistency with the NMR parameters. Twenty-three structures (Table 4) showed consistency with the NMR-derived H-bonding pattern, had E_{tot} within 60 kcal mol^{-1} of the structure with the lowest E_{tot} and $< r_{viol} >$ less than 0.2 Å. RMD calculations [20–22] were carried out on these selected structures.

After RMD calculations the average backbone RMS for the whole peptide was 5 ± 1 Å; it decreased to 3.2 ± 0.8 Å when the region 2–10 was considered and to 2.2 ± 0.7 Å in the case of region 11–19.

As with peptide A, a conformation was assigned to each residue of the 23 structures using the Ramachandran classification (Table 5). This led to structures 1, 2, 3, 5, 11, 13 and 20 being considered as less likely to represent the solution structure of the peptide, since one or more than one residue in each of these structures was found well-outside the allowed regions of the Ramachandran plot (Table 5). Structure 9 was discarded because of a *cis* amide bond between R5 and G6. Structure 17, which had the shape of a large cycle with the N-terminal region embedded in the cycle and within H-bonding distances of the C-terminal residues, was considered unlikely to occur in the virus structure which contains this peptide. The 14 structures thus remaining were then analysed and the structural features of individual regions of the peptide discussed hereafter.

The C-terminal Region. Inspection of Table 5 as well as computer superposition (Figure 6) revealed that the 14 structures were less homogeneous than those found for peptide A. However, as with the latter, the C-terminal region (sequence 11–17) was essentially the same in all examples and mainly helical (Figure 7). Proline was inside the first turn of this helix at position $i+3$. This induced A12 to adopt almost exclusively an A(αr) (A hereafter) conformation (see Table 5 for definition of A, B, etc.) instead of the B(β) (B hereafter) conformation generally found in Ala-Pro dipeptides [23].

The H-bonding pattern around P13 (Table 6 and Figure 8) and the average backbone angles at P13 (Table 7) excluded the existence of a γ-turn. This was not in contrast with the NMR data since the NH of R14 was, in all structures, shielded from the solvent.

The Central Region of the Peptide. H-bonds between COOH–D7 and either the OH–S8 and NH–S8 (structures 4, 6, 8, 14, 15, 18, 19 and 21) (Table 6) or CO–S10 (structure 23) concentrated the side chain of D7 in a limited conformational space (Figure 9). A similar conclusion applied to structures 12, 16 and 22 where the H-bond was between OH–S10 and CO–G6. These H-bonds generated a loop comprising residues 7 to 10 (Figures 10 and 11). F8 occupied one of the corners of the loop and had its aromatic ring pointing outward and solvent exposed. The three residues inside the loop, i.e. F8, G9 and S10, were assigned to a short extended strand which contrasted with the helix assignment based on a qualitative interpretation of the NOE data (see above). NH–S10 was either free but solvent-shielded or H-bonded to COOH–D7 (Table 6), consistent with its small NMR temperature coefficient.

The above conclusions applied to nearly all the structures investigated. The only exceptions were

Table 5 Ramachandran Classification of Each of the Residues Contained in the 23 Structures of Peptide USA Obtained after DG and RMD Calculations

	Residue *n*																	
Structure *n*	S2	G3	V4	R5	G6	D7	F8	G9	S10	L11	A12	P13	R14	V15	A16	R17	Q18	L19
USA1	***L***	***t***	***L***	***B***	***t***	***B***	***A***	***t***	***B***	***A***	***A***	***H***	***A***	***A***	***A***	***A***	***L***	***B***
USA2	***B***	***t‴***	***L***	***B***	***t***	***B***	***A***	***t″***	***A***	***H***	***A***	***A***	***A***	***A***	***A***	***A***	***L***	***B***
USA3	***L***	***t′***	***B***	***B***	***t‴***	***A***	***H***	***t″***	***B***	***A***	***A***	***B***	***B***	***L***	***A***	***A***	***H***	***B***
USA4	L	t‴	A	B	t	B	A	t″	B	A	A	A	A	A	A	A	L	B
USA5	***L***	***t‴***	***B***	***B***	***t***	***B***	***H***	***t′***	***H***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***B***
USA6	L	t‴	B	B	t	B	A	t	A	A	A	A	A	A	A	A	A	B
USA7	L	t‴	B	B	t‴	B	A	t′	A	A	A	A	A	A	A	A	B	L
USA8	L	t′	L	L	t‴	B	A	t	B	A	A	A	A	A	A	A	A	B
USA9	***L***	***t‴***	***L***	***B***	***t″***	***B***	***A***	***t″***	***B***	***A***	***A***	***A***	***A***	***A***	***B***	***L***	***A***	***B***
USA10	B	t″	L	B	t	B	A	t″	B	A	A	A	A	A	B	L	L	B
USA11	***B***	***t‴***	***A***	***B***	***t′***	***H***	***A***	***A***	***A***	***A***	***A***	***B***	***L***	***L***	***B***	***L***	***L***	***B***
USA12	L	t‴	B	B	t	A	A	t′	B	A	A	A	A	A	A	A	B	B
USA13	***L***	***t‴***	***b***	***b***	***T***	***B***	***A***	***t***	***A***	***H***	***A***	***A***	***A***	***A***	***A***	***B***	***A***	***B***
USA14	B	t″	B	B	t	B	A	t′	B	A	A	A	A	A	A	A	L	B
USA15	L	t‴	L	L	t	A	L	t′	A	A	B	A	A	A	A	A	B	B
UA17	***A***	***t′***	***B***	***B***	***t***	***B***	***A***	***t′***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***B***	***L***
USA18	L	t″	B	B	t″	B	B	t″	A	A	A	A	A	A	A	A	A	B
USA19	L	t″	B	B	t‴	B	A	t′	A	A	A	A	A	A	A	A	A	B
USA20	***B***	***t‴***	***B***	***B***	***t***	***B***	***A***	***t‴***	***H***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***L***
USA21	A	t‴	A	B	t	B	A	t‴	H	A	A	A	A	A	A	A	L	L
USA22	B	t‴	B	B	t	A	A	t″	B	A	A	A	A	A	A	A	B	B
USA23	L	t‴	L	B	t‴	B	A	t″	B	A	A	A	A	A	B	L	B	B

The Ramachandran classifications were expressed as H = disallowed regions, A = α-helix, B = β-sheet. L = left-handed helix and for Gly residues as t, t′, t″ e t‴ where the latter refer to the upper left, upper right lower left and lower right quadrant respectively of the ϕ, φ distribution map.

Each classification, which contained the most favoured regions of the Ramachandran plot as well as additionally allowed regions, was carried out using the program PROCHECK (Lasowsky *et al.*, 1993).

The first (Gly) and last (Pro) residues of the sequence are not shown since their NMR parameters were not sufficiently defined.

Those structures which, owing to residues in disallowed region of the ϕ, φ distribution map, were not considered in the discussion (see text) are shown in bold and italic.

structure 7 where the loop region was completely absent and structure 10 where it was less pronounced.

The RGD Region. The RGD regions of the 14 structures are superimposed in Figure 10. Turns stabilized by H-bonds were found only in a few cases (e.g. inverse γ-turn NHG6–COV4 – structures 10, 12 and 14 – β II turn NHD7–COV4 – structure 21). The remaining models did not display any H-bond involving the RGD triplet and their backbone dihedral angles were not consistent with regular turns. The NHs of G6 and D7 were shielded from the solvent in agreement with NMR temperature studies (Table 3). The conformation of R5 and D7 was essentially the same in all 14 structures (Table 5).

DISCUSSION

NMR parameters were derived for a peptide corresponding to the immunodominant region (141–160) of FMDV serotype A, variant USA. Although the majority of these parameters indicated conformational averaging, amide proton temperature coefficients, which ranged from 1.9 to 7.4, suggested the existence of H-bonded structures. This conclusion was in agreement with previous CD studies conducted in TFE [2, 3] where the spectrum of peptide USA was characterized by minima at 206 and 222 nm. Since the main aim of this work was to assess the structural differences between peptides A and USA, NOE constraints were used to generate 3D hypotheses which could explain the different serological behaviour of these two peptides.

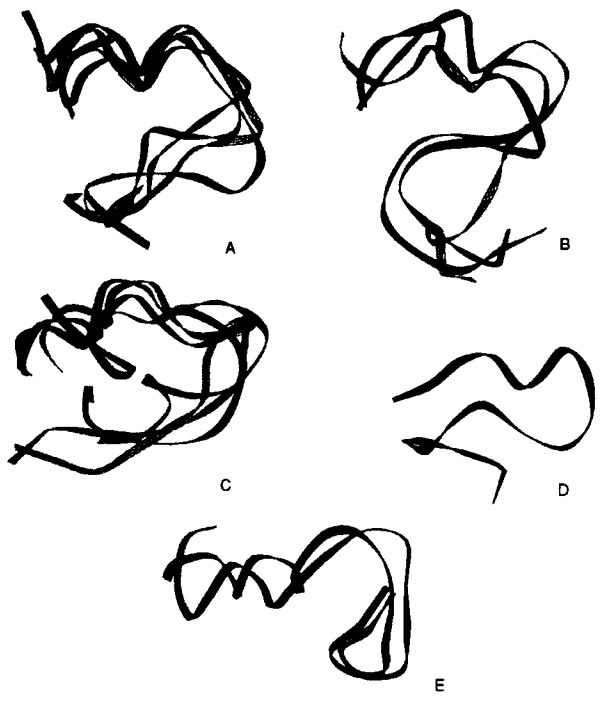


Figure 6 Ribbon structure of the 14 models of peptide USA obtained after DG, EM and RMD calculations. The structures are grouped according to their similarities calculated over the V4–Q18 region. The N- and C-terminal residues were not well-defined owing to a small number of NOE constraints obtained both from 2D- and 3D-NOESY spectra. (A) structures 4, 12, 16, 22; (B) structures 7, 10, 23; (C) structures 6, 8, 14, 18; (D) structures 15; (E) structures 19 and 21.

The models obtained were less homogeneous than those derived for peptide A. However, when the various regions of the peptide were analysed separately, several general characteristics were identified which were common to the majority of the models and consistent with both other NMR derived parameters and previous CD studies. In particular three separate regions could be identified which had the following characteristics:

1. A helical region spanning the entire C-terminal sequence of the peptide and containing a proline residue in the less common position $i+3$ of the first turn. A further characteristic of the helix was the conformation of A12 which was of type A in contrast to the expected B conformation predicted for Ala-Pro dipeptides [23].
2. A loop involving residues 7 to 10 which was common to the majority of the structures. This turn was stabilized by H-bonds involving either residues D7 and S10 or G6 and S10. The position of the loop (i.e. adjacent to the C-terminal helix) and the type of residues (i.e. D7 and S10) involved in its stabilization were reminiscent of the 'capping box' [24, 25]. The latter is defined as the helix stop signal region of peptides and proteins and is generally characterized by the presence of a reciprocal backbone–side chain H-bonding interaction between residue i and $i+3$ [25]. In the case of peptide USA the H-bonding interaction was found mainly between D7 (residue i of the 'capping box') and the OH of S10 ($i+3$) and not with the NH of the latter as instead described for those proteins where the capping box has been identified. However, the average ϕ and ψ angles of D7 in the majority of structures were $-92° \pm 38°$ and $100° \pm 40°$ and compared favourably with those expected for the first residue of the capping box [25].
3. The RGDF sequence which was characterized by a turn around G6 and, in the majority of cases, not stabilized by H-bonds. The conformation of residues R5, D7 and F8 were mostly constant in the structures investigated and different from those of equivalent residues in peptide A.

The amino acid substitution at position 8 (148 in the virus structure) has been proposed to play a key role in the serological tests involving the FMDV peptides. Furthermore this position is adjacent to the RGD triplet, the cell binding site which has been shown to partially overlap with the peptide/virus antibody binding site [26]. Any structural difference found at the level of position 8 and adjacent residues might therefore account for the different serological behaviour of peptides A and USA.

Indeed comparison of the Ramachandran conformation indicated that the largest levels of difference were in the N-terminal half of the peptides, the C-terminal region being similar and largely helical in both cases. Within the RGDX (X = S, F) sequence the conformation of residues R5, D7 and F8 was essentially the same in all models and different fron that of the same residues in peptide A (Table 5 and accompanying paper).

The influence of residue 8 on the adjacent region is illustrated in Figure 11 and discussed hereafter. In peptide A, where the side chains of D7 and S8 were within H-bond distance, an additional H-bond, that between NH–D7 and CO–A12, generated a large loop whose corners were occupied by S8 and S10 (Figure 12). In peptide USA, where a H-bond between D7 and

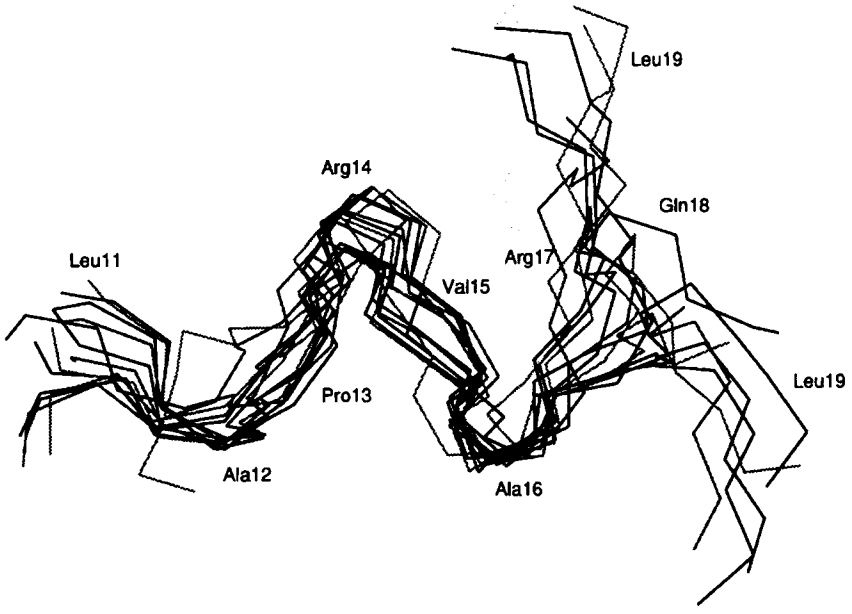


Figure 7 The 11–19 region of 14 structures of peptide USA obtained from computer calculations and a screening procedure which took into account the residues ϕ and Ψ distribution map. The figure illustrates the helical character of the C-terminal region which in some of the structures (see Table 5) was interrupted at the level of Q18.

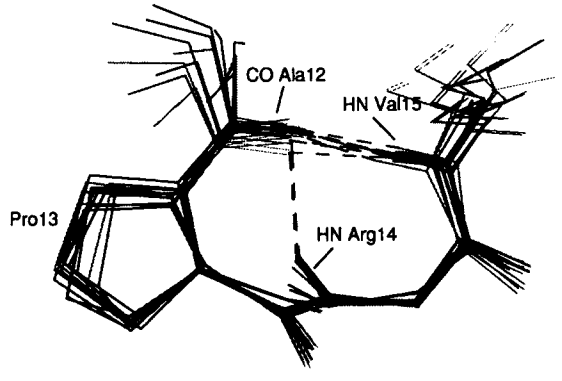


Figure 8 Superposition of the backbone of the 12–15 region of 14 structure of peptide USA. The H-bonds between CO12 and either NH14 or NH15 are shown.

F8 was not possible, the turn was restricted to fewer residues and stabilized by interactions between S10 and either D7 or G6 (structure 7 which in this region was in an extended conformation was the only exception to these general characteristics).

As to any perturbation arising from the second amino acid substitution, that at position 13, there were contrasting arguments. Thus in peptide A CO–A12 was H-bonded to NH–D7 whereas in peptide USA it formed H-bonds with residues C-carboxyl to P13. Therefore, based on these observations only, the substitution at position 13 would have a profound impact on the structure of RGDX and flanking residues.

On the other hand the chemical nature of the side chains of residues 7 and 8 could be indirectly responsible for the long-range interactions found in both peptides, by favouring in one case (peptide A) the H-bond between COOH–D7 and OH–S8 while forcing, in the other case, the side chains of D7 and F8 to adopt an anti-parallel orientation. The COOH group would then be free to form a H-bond with OH–S10. In this assumption the structure of the N-terminal half of the peptide would be independent of the mutation at position 13. The data presented here were not sufficient to distinguish between the two hypotheses.

Acknowledgements

This work has been supported by the Wellcome Trust (grant no. 18937/1.5). L. Z. and H. M. have been supported by the 'Progetto finalizzato Chimica fine e secondaria' of the Italian CNR (sottoprogetto 3.2). L.

Table 6 List of H-bonds Present in the Selected Structures[a]

H-bond	Structures												
NG(G_6)–CO(V_4)	8	10	12	14									
NH(S_8)–$^{\delta}$CO(D_7)	6	8	12	14	16								
NH(R_{14})–CO(A_{12})	7	8	10	14	23								
NR(V_{15})–CO(A_{12})	4	6	7	8	10	12	14	15	16	18	21	22	23
NH(A_{16})–CO(A_{12})	4	7	10	12	15	16	18	19	21	22			
NH(Q_{18})–CO(V_{15})	4	6	7	8	15	21							
NH(L_{19})–CO(R_{17})	7	12	15	16	23								
NH(S_{10})–$^{\delta}$CO(D_7)	4	6	8	14	15	19							
NH(L_{11})–$^{\gamma}$O(S_{10})	6	10	16	18	21								
$^{\epsilon}$NH(R_{14})–CO(L_{11})	6	14	18	22	23								
$^{\epsilon}$NH(R_6)–$^{\delta}$CO(D_7)	4	6	10	12									
NH_2(R_{14})–CO(S_{10})	7	10	19	21	23								

[a] Only those H-bonds present in at least 4 of the 14 selected structures (4, 6, 7, 8, 10, 12, 14, 15, 16, 18, 19, 21, 22, 23) are reported.

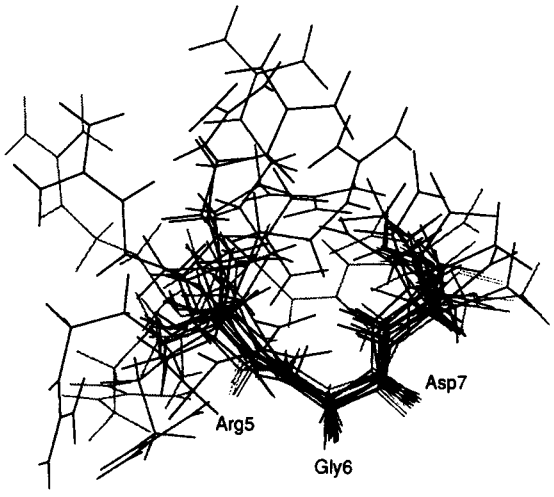


Figure 9 Backbone superposition at the RGD triplet of the 14 structures of peptide USA considered in the text.

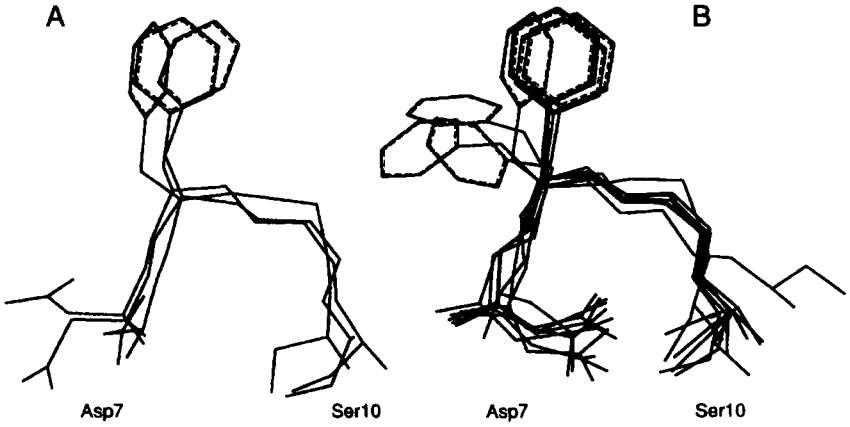
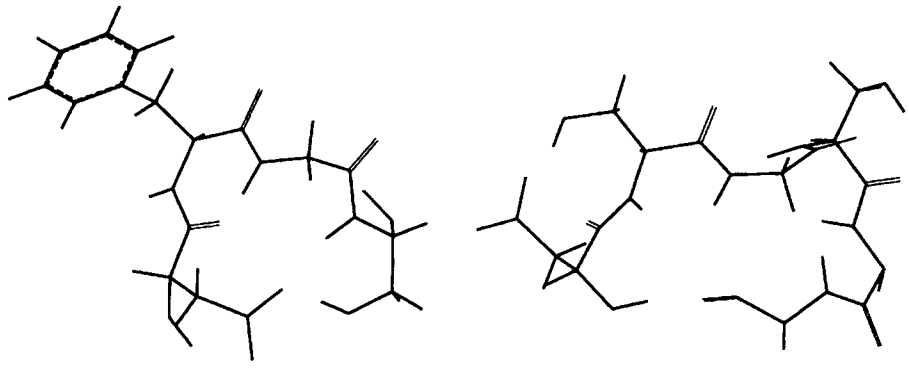


Figure 10 The D7–S10 region of peptide USA forms a loop stabilized by H-bonds between either residues G6 and S10 (A, structures 12, 16 and 22) or residues D7 and S10 (B, structures 4, 6, 7, 10, 14, 18, 19 and 21).

Table 7 Average ϕ and Ψ Dihedral Angles of Residues V4 and G9–Q18 Contained in the 14 Structures Discussed in the Text

	Dihedral Angle	V4[b]	G97sup;b,c	S10[b]	L11[b]	A12[b]	P13[b,d]	R14[b]	V15[b]	A16[b]	R17[b]	Q18[b]
+ + +[a]	ϕ	−106 ± 26	149 ± 44	−110 ± 32	−104 ± 29	−64 ± 5	−47 ± 11	−54 ± 18	−82 ± 19	−61 ± 7	−78 ± 13	−87 ± 7
	ψ	76 ± 8	75 ± 20	−170 ± 33	−67 ± 23	−43 ± 7	−18 ± 14	−50 ± 10	−32 ± 14	−40 ± 16	−22 ± 1	63 ± 8
+ +[a]	ϕ	57 ± 17	−128 ± 49	−96 ± 35		−66				−78 ± 2	55 ± 6	−70 ± 7
	ψ	78 ± 15	−70 ± 24	−71pm;29		151				94 ± 12	84 ± 5	−48 ± 11
+[a]	ϕ	−107 ± 25									−152	67 ± 17
	ψ	−70 ± 12									73	71 ± 13

[a] For each residue, a maximum of three different sets of dihedral angles were obtained from the Ramachandran classification of Table 5. For instance R17 was found to be in either A (11 cases), L (2 cases) or B conformation (1 case). Thus, for R17, (+ + +) refers to the average dihedral angles of all the structures with a A conformation, (+ +) to the L conformation and (+) to the B conformation.
[b] Average dihedral angles are shown with their SD.
[c] In the case of Gly-9 the angles of structures 6 and 8 were computed together with those of class (+ + +) since their ϕ angles were very close to −180° and the Ψ angles all positive.
[d] In the case of proline only 12 structures were considered. The ϕ and Ψ angles of structure 6 were −2° and −67° respectively, those of structure 18, −13° and −59°.



Str. USA 4 Str. A 7

Figure 11 Comparison between the lop region of two representative models of peptide USA and peptide A. In the case of peptide USA, structure 4 shows the loop between residues D7 and S10 stabilized by a H-bond involving the side chains of these two residues. In the case of structure 7 of peptide A the side chain of D7 forms a H-bond with the CO of A12 thus establishing a larger loop between these two residues.

Z. and H. M. are indebted to Italfarmaco for economical support, Drs A. L. Segre and E. Ragg for technical help on the 600 MHz instrument and F. Greco for technical assistance.

REFERENCES

1. D. J. Rowlands, B. E. Clark, A. R. Carrol, F. Brown, B. H. Nicholson, J. L. Bittle, R. A. Houghten and R. A. Lerner (1983). Chemical basis of antigenic variation in foot-and-mouth disease virus. *Nature, 306*, 694–697.
2. G. Siligardi, A. F. Drake, P. Mascagni, D. Rowlands, F. Brown and W. A. Gibbons (1991). A CD strategy for the study of polypeptide folding/unfolding. A synthetic foot-and-mouth disease virus immunogenic peptide. *Int. J Peptide Protein Res., 38*, 519–527.
3. G. Siligardi, A. F. Drake, P. Mascagni, D. Rowlands, F. Brown and W. A. Gibbons (1991). Correlation between the conformations elucidated by CD spectroscopy and the antigenic properties of four peptides of the foot-and-mouth disease virus. *Eur. J. Biochem., 199*, 545–551.
4. G. W. Vuister, R. Boelens, A. Padilla, G. J. Kleywegt and R. Kaptein (1990). Assignment strategies in homo-nuclear three-dimensional ^{1}H NMR spectra of proteins. *Biochemistry, 29*, 1829–1834.
5. A. Bax, M. Ikura, L. E. Kay, D. E. Torchia and R. Tschudin (1990). Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Reson., 86*, 304–309.
6. H. Molinari, G. Esposito, R. Consonni, M. Pegna and L. Zetta (1992). Distance evaluation via heteronuclear SQC-NOESY experiments. *J. Biomol. NMR*, 289–299.
7. K. Wuethrich: *NMR of Proteins and Nucleic Acids*, Wiley, New York 1986.
8. C. Griesinger, G. Otting, K. Wuethrich and R. R. Ernst (1988). Clean TOCSY for ^{1}H spin system identification in macromolecules. *J. Am. Chem. Soc., 110*, 7870–7872.
9. S. Macura and R. R. Ernst (1980). Elucidation of cross relaxation in liquids by two-dimensional NMR. *Mol. Phys., 41*, 95–117.
10. D. Neuhaus and M. Williamson: *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, New York 1989.
11. E. T. Olejniczak, R. T. Gambe Jr and S. w. Fesik (1986). Accounting for spin diffusion in the analysis of 2D NOE data. *J. Magn. Reson., 67*, 28–32.
12. G. Esposito and A. Pastore (1988). An alternative method for distance evaluation from NOESY spectra. *J. Magn. Reson., 76*, 331–336.
13. G. C. Johnson Jr, T. G. Pagano, C. T. Basson, J. A. Madri, P. Gooley and I. M. Armitage (1993). Biologically active Arg-Gly-Asp oligopeptides assume a type II β-turn in solution. *Biochemistry, 32*, 268–273.
14. M. C. Vega, C. Aleman, E. Giralt and J. J. Perez (1992). Conformational study of a nine residue fragment of the antigenic loop of foot-and-mouth disease virus. *J. Biomol. Struct. Dynamics, 10*, 1–8.
15. G. D. Rose, L. M. Gierasch and J. A. Smith (1985). Turns in peptides and proteins. *Adv. Protein Chem., 37*, 1–109.
16. T. F. Havel, J. D. Kuntz and G. M. Crippen (1983). The theory and practice of distance geometry. *Bull. Math. Biol 45*, 655–720.
17. T. F. Havel and K. Wütrich (1984). A distance geometry program for determining the structures of small proteins and other macromolecules from nuclear magnetic resonance measurements of intramolecular ^{1}H–^{1}H proximities in solution. *Bull. Math. Biol., 46*, 673–698.

18. R. Fletcher and C. M. Reeves (1964). Function minimization by conjugate gradients. *Comput. J.*, *7*, 149–167.
19. W. Braun (1987). Distance geometry and related methods for protein structure determination from NMR data. *Q. Rev. Biophys.*, *19*, 115–157.
20. G. M .Clore, A. M. Gronenborg, A. T. Brünger and M. Karplus (1985). Solution conformation of a heptadecapeptide comprising the DNA binding helix F of the cyclic AMP receptor protein of *Escherichia coli*. Combined use of ^{1}H nuclear magnetic resonance and restrained molecular dynamics. *J. Mol. Biol.*, *186*, 435–455.
21. R. Kaptein, E. R. P. Zuiderberg, R. M. Scheek, R. Boelens and W. F. van Gunsteren (1985). A protein structure from nuclear magnetic resonance data of lac repressor headpiece. *J. Mol. Biol.*, *182*, 179–182.
22. J. de Vlieg, R. M. Scheek, W. F. van Gunsteren, H. J. C. Berendsen, R. Kaptein and J. Thomason (1988). Combined procedure of distance geometry and restrained molecular dynamics techniques for protein structure determination from nuclear magnetic resonance data: application to the DNA binding domain of lac repressor from *Escherichia coli*. *Proteins: Struct. Funct. Genet.*, *3*, 209–218.
23. C. R. Cantor and P. R. Schimmel: *Biophysical Chemistry. Part I: The Conformation of Biological Macromolecules*, W. H. Freeman, New York 1980.
24. E. N. Baker and R. E. Hubbard (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.*, *44*, 97–179.
25. E. T. Harper and G. D. Rose (1993). Helix stop signals in proteins and peptides: the capping box. *Biochemistry*, *32*, 7605–7609.
26. G. Fox, N. R. Parry, P. V. Barnett, B. McGinn, D. J. Rowlands and F. Brown (1989). The cell attachment site on foot-and-mouth disease virus includes the amino acid sequence RGD (arginine–glycine–aspartic acid). *J. Gen. Virol.*, *70*, 625–637.