Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules

Kirsten Falk, Olaf Rötzschke, Stefan Stevanović*, Günther Jung* & Hans-Georg Rammensee†

Max-Planck-Institut für Biologie, Abteilung Immungenetik, Corrensstrasse 42, W-7400 Tübingen, Germany
* Institut für Organische Chemie, Universität Tübingen, Auf der Morgenstelle 18, W-7400 Tübingen, Germany

The crystal structures of major histocompatibility complex (MHC) molecules contain a groove occupied by heterogeneous material thought to represent peptides central to immune recognition, although until now relatively little characterization of the peptides has been possible. Exact information about the contents of MHC grooves is now provided. Moreover, each MHC class I allele has its individual rules to which peptides presented in the groove adhere.

T LYMPHOCYTES recognize their antigens in context of MHC-encoded molecules, a phenomenon called MHC restriction[1-5]. Crystallography of human MHC class I molecules, HLA-A2 and Aw68, revealed a groove made up by the $\alpha 1$ and $\alpha 2$ domains of heavy chains[3,6]. This groove is believed to be the binding site for antigenic peptides, as both crystals contain peptide-sized structures not compatible with the MHC sequences located at that groove[6]. T cells can recognize synthetic peptides loaded on MHC class I molecules, and MHC-associated peptides representing T-cell epitopes have been extracted from normal or virus-infected cells[2,4,5,7,8]. Similarly, the antigens recognized by MHC class II-restricted T cells can be mimicked by artificial peptides[9], and MHC-associated antigenic peptides have been eluted from MHC class II molecules[10]. By virtue of their position in the middle of trimolecular complexes made up of T-cell receptor, peptide and MHC molecule[11], T-cell epitopes are literally central to the specific immune system and the so far unknown rules defining them should be central to understanding it[12-15].

Comparison of the first naturally processed H-2K^d-restricted T-cell epitope[4] with several peptides that contain other K^d-restricted epitopes revealed some common features not found in non-K^d-restricted epitopes[4,7]. Assuming that all naturally processed K^d-restricted epitopes are nonapeptides, all those peptides containing K^d-restricted epitopes could be aligned with a Tyr residue at the second position and an amino-acid residue with a side-chain methyl group (Val, Ile, Thr, Ala or Leu) at the last position, suggesting a K^d-specific peptide motif. (Tyr is important in K^d-restricted peptides, and some allele-specificity of T-cell epitopes has been noticed before[12-16].) Sequencing the self-peptide blends eluted from purified K^d molecules should reveal this K^d-specific motif. If so, sequencing of peptide mixtures from other MHC class I molecules should reveal the respective motifs if they exist.

Elution of peptides from K^d molecules

K^d molecules were immunoprecipitated from detergent extracts of P815 tumour cells (H-2^d). Bound peptides were dissociated from bead-associated K^d molecules by acid treatment[7] and separated by reversed-phase HPLC (Fig. 1*a*, *b*). Precipitation from influenza-infected cells (not shown) gave a K^d-restricted influenza epitope in fraction 24, indicating that the method can isolate MHC-bound peptides. Heterogeneous material elutes in low amounts between fractions 20 and 28 (Fig. 1*a*, *b*), covering the range of many MHC class I-restricted T-cell epitopes with this HPLC-gradient (refs 4, 7, 8, and O.R. *et al.*, manuscript submitted).

K^d-restricted peptide motif

Fractions 20–28 were pooled both from the K^d-derived batch and from a mock precipitate, and both batches sequenced automatically using the Edman degradation method (Table 1). This method involves the sequential derivatization and removal of amino acids from the N terminus, each of which is identified chromatographically. Because it is unusual to sequence complex mixtures of peptides, we have presented the raw data from the sequencer. Table 1*a* and *b* shows the results from two sequencing attempts for K^d-eluted peptides. Table 1*c* shows the sequencing result of a mock elution with D^b-specific antibodies on P815 lysates. The K^d-eluted peptides have a distinct amino-acid residue pattern for each position from 1 to 9, whereas the mock-eluted material shows a uniform pattern of residues throughout, with a decrease of the absolute amount of each residue with every cycle. Thus, for the K^d-eluted peptides, only residues showing more than 50% increase in the absolute amount compared with the previous or the pre-previous cycle were arbitrarily considered significant and are underlined. The first position is difficult to judge; there is no previous cycle, and all free amino acids present in the HPLC pool are detected in this position. For the second position, the only residue whose frequency is clearly increased by comparison with the previous cycle is Tyr, in both attempts (Table 1*a*, *b*), for example, 60.9 to 875.6 pmol. The only other residue that shows an increase, however marginal, is Phe, which has a side chain similar to that of Tyr. This confirms our premises resulting from comparing the natural K^d-restricted influenza epitope TYQRTRALV (single-letter amino-acid code) with other K^d-restricted peptides, with regard to the Tyr at position 2. By contrast to the second position, there is no single amino-acid residue standing out in the following positions up to 8, although up to 14 different residues are detected in the individual positions. At position 9, the residues detected are Ile and Leu. This again agrees with our premises. There is no signal increase at position 10, indicating that most K^d-bound self-peptides are no longer than 9 residues. The natural K^d-restricted influenza peptide is also a nonapeptide[4]. The consensus sequence pattern indicated by these data is shown in Table 1*d*. Most evident are Tyr at position 2 and Ile or Leu at 9, whereas at all other positions a larger, but distinct set of residues was found. A comparison of this motif with peptide sequences containing K^d-restricted epitopes indicates that, if aligned by their Tyr residues, most fit quite well with the consensus K^d-restricted nonamer motif (Table 1*d*).

Sequence of a prominent self-peptide

The peak marked by an arrow in fraction 29 of Fig. 1*b* and the corresponding fraction of the mock precipitation were rechromatographed giving higher resolution (Fig. 1*c*). The sharp specific peak was determined to be SYFPEITHI by direct

† To whom correspondence should be addressed.

TABLE 1 Sequencing of the self-peptide mixture eluted from immunoprecipitated K^d molecules

Cycle	A Ala	R Arg	N Asn	D Asp	E Glu	Q Gln	G Gly	H His	I Ile	L Leu	K Lys	M Met	F Phe	P Pro	S Ser	T Thr	Y Tyr	V Val
(*a*) Experiment 1	Amino-acid residues (in pmol)																	
1	172.8	46.1	44.9	13.6	73.5	317.8	171.6	3.2	73.1	66.5	231.2	28.0	35.3	56.7	145.2	73.3	60.9	130.9
2	25.6	14.1	10.1	7.7	18.7	71.9	71.9	1.2	28.4	22.6	13.9	11.1	97.7	14.8	14.6	9.3	875.6	18.8
3	88.7	26.7	51.5	10.0	23.1	86.8	62.5	2.9	183.2	308.7	71.6	25.6	41.5	13.5	24.0	22.0	66.1	150.2
4	158.5	14.2	31.9	17.9	53.3	44.8	85.2	6.7	32.1	36.6	29.5	9.2	5.8	226.9	26.2	19.9	14.7	41.5
5	139.0	30.1	42.2	22.9	15.1	44.1	154.5	1.8	59.3	86.6	10.2	50.8	2.6	87.8	64.2	47.6	8.8	104.2
6	116.5	29.2	42.6	13.0	10.6	38.3	139.1	8.5	90.1	99.9	194.5	69.7	27.5	38.6	15.1	26.5	35.9	106.8
7	51.5	79.7	125.1	25.8	47.0	73.7	65.8	7.9	12.8	23.4	37.8	11.2	5.1	16.9	39.3	148.4	11.2	36.1
8	44.2	29.0	48.9	22.4	75.8	58.0	59.0	18.3	10.1	30.4	41.5	10.5	19.3	10.8	28.8	46.0	47.9	63.2
9	13.0	8.3	20.1	10.7	14.4	10.4	20.5	3.5	129.4	155.2	3.9	4.9	5.0	7.2	7.0	10.1	9.4	35.4
10	6.5	4.4	7.8	6.1	4.2	5.6	14.6	1.3	32.1	58.3	3.1	1.8	3.1	4.7	4.2	5.2	4.3	8.8
(*b*) Experiment 2																		
1	54.5	0.4	5.8	3.5	5.0	5.8	62.5	1.8	11.2	13.2	35.3	5.8	11.5	35.3	57.8	26.0	15.1	29.2
2	14.1	0.2	1.2	1.0	2.2	3.6	20.0	0.5	3.4	5.7	3.4	1.6	19.6	8.6	8.5	5.1	187.7	5.5
3	22.4	4.4	10.3	2.5	7.1	15.9	26.2	0.8	41.0	77.2	12.7	7.5	23.0	6.6	6.7	5.3	16.9	22.7
4	48.2	1.4	11.7	5.5	13.8	8.1	34.3	2.3	7.3	10.4	4.9	3.7	2.1	60.0	6.9	5.7	3.8	12.1
5	35.2	1.7	11.7	8.0	9.1	7.2	41.5	0.7	12.3	18.1	1.4	17.6	0.9	20.7	16.1	11.6	1.7	25.6
6	32.3	5.4	7.9	5.0	6.4	6.5	35.9	1.8	32.4	31.9	31.4	19.9	4.5	0.4	4.2	3.5	5.5	27.8
7	11.2	1.1	27.7	11.8	17.2	15.7	16.0	2.7	5.7	7.0	5.9	2.9	1.1	1.5	12.4	47.3	2.0	9.0
8	10.7	3.4	7.8	7.3	16.5	9.7	19.5	4.3	2.5	8.7	5.0	2.4	4.2	0.8	7.6	10.7	8.2	16.8
9	4.1	2.6	4.0	4.2	4.8	1.9	10.6	0.4	37.0	26.6	0.0	1.3	1.5	0.5	2.3	3.1	1.8	7.7
10	2.5	1.0	1.3	3.1	2.7	1.0	7.5	0.2	13.0	13.5	0.0	1.0	1.3	1.5	1.6	1.4	1.2	3.4
(*c*) Sequencing of mock-precipitated material*																		
1	63.5	5.6	3.6	3.9	8.3	11.3	51.5	2.3	12.2	16.5	8.4	3.5	10.8	47.0	35.2	27.3	12.7	24.4
2	24.8	2.5	3.1	3.6	7.9	6.2	33.8	1.3	6.9	12.1	4.5	1.4	5.8	18.4	7.4	6.4	6.9	13.8
3	15.2	0.9	2.5	3.0	6.6	3.6	26.6	1.2	4.1	11.0	2.7	1.2	4.2	16.1	2.7	4.0	4.3	8.6
4	11.5	1.0	2.2	3.2	5.7	2.6	19.5	0.8	3.9	7.3	2.8	1.1	2.7	10.7	1.6	2.4	3.1	6.4
5	10.5	1.4	2.1	3.1	5.0	2.6	15.7	1.0	3.1	6.2	2.3	0.7	2.2	7.9	0.9	1.7	2.6	5.2
6	8.8	1.1	1.6	3.1	4.1	2.0	12.6	1.1	2.2	4.6	1.9	0.6	1.9	6.5	1.1	1.4	1.9	3.9
7	6.8	1.0	1.6	2.4	3.5	1.8	9.8	0.5	1.8	3.4	2.1	0.4	1.7	4.3	1.6	1.5	1.7	2.7
8	0.0	0.3	0.0	2.1	0.2	0.8	0.8	0.6	1.1	2.8	1.7	0.3	1.1	3.6	0.9	2.2	0.2	2.6
9	0.1	0.6	0.0	1.8	0.0	0.8	0.7	0.2	1.6	2.5	1.7	0.5	1.1	3.3	1.3	1.7	0.1	2.1
10	0.2	0.3	0.0	1.7	0.1	0.5	0.8	0.2	1.0	2.5	1.4	0.3	1.3	2.7	0.8	1.7	0.1	2.1

(*d*) The K^d-restricted peptide motif†

	Position 1	2	3	4	5	6	7	8	9‡
Dominant anchor residues		**Y**							**I L**
Strong			**N I L**	**P**	**M**	**K F**	**T N**		
Weak	K A R S V T	F	A H V R S F E Q K M T	A E S D H N	V N D I L S T G	H I M Y V R L	P H D E Q S	H E K V V F R	

Known epitopes, aligned§										Protein source	Ref.
	T	Y	Q	R	T	R	A	L	V	Influenza PR8 NP 147–154	4, 29
	S	Y	F	P	E	I	T	H	*i*	Self-peptide of P815	This paper
	I	Y	A	T	V	A	G	S	L	Influenza JAP HA 523-549‖	30, 31
	V	Y	Q	I	L	A	I	Y	A	Influenza JAP HA 523-549‖	30, 31
	I	Y	S	T	V	A	S	S	L	Influenza PR8 HA 518–528	32
	L	Y	Q	N	V	G	T	Y	V	Influenza JAP HA 202–221	30, 31
	R	Y	L	E	N	G	K	E	T	*L*¶ HLA-A24 170-18233	33
	R	Y	L	K	N	G	K	E	T	*L*¶ HLA-Cw3 170-186	34
	K	Y	Q	A	V	T	T	T	L	P815 tumour antigen	35
	S	Y	I	P	S	A	E	K	I	*Plasmodium berghei* CSP 249-260	36
	S	Y	V	P	S	A	E	Q	I	*Plasmodium yoelii* CSP 276-288	37

Fraction 20–28 of K^d-precipitates (*a*, *b*) or mock precipitate (*c*) prepared from $10–20 \times 10^9$ P815 cells according to the precipitates in Fig. 1*a*, *b* were pooled and sequenced by Edman degradation, performed in a pulsed-liquid protein sequencer 477A equipped with an on-line PTH-amino-acid analyser 120A (Applied Biosystems). Glass-fibre filters were coated with 1 mg BioBrene Plus (Applied Biosystems) and were not precycled. Sequencing was carried out using the standard programmes BEGIN-1 and NORMAL-1 (Applied Biosystems). Cys was not modified and therefore not detectable, PTH-Trp coeluted with diphenylurea and, in three out of six experiments, PTH-Arg coeluted with the major derivative of PTH-Thr. Thus, Cys and Trp are not detectable, and Arg only in the absence of Thr. The numbers indicate pmol of individual amino-acid residues detected at each cycle. Numbers indicating more than 50% increase in absolute amount of individual residues compared with their amount in either of the two previous cycles (but not if a value decreases compared with the previous position, or if a value is below 1 pmol) are underlined. These are the residues arbitrarily considered significant at the respective position. At least 10% of a signal is caused by a lag effect from the previous cycle, in addition 10% of Asn is hydrolysed to Asp and 20% of Gln to Glu. Increases in Arg signals can be due to a hydrolysed derivative corresponding to the major derivative of Thr. Thus, increases of Asp in presence of strong Asn signals are not significant; the same is true for Gln–Glu and Arg–Thr pairings, respectively. Material extracted from glycine beads (see Fig. 1) did not indicate particular amino acids except for some glycine residues (not shown).

* The data for material eluted from anti K^b-beads used to mock-precipitate P815 lysates. Sequencing of material from P815 lysates mock-precipitated with anti-D^b-beads showed a very similar pattern, with no increase of any residue compared with its amount at a previous cycle.

† Residues at individual positions were classified into those dominant within anchor positions, or for those with strong or weak signals, according to the magnitude of increase seen for the respective residue.

‡ Anchor positions are indicated in bold types. Classification 'anchor' is only applied if a position reveals a strong signal for only one residue, or alternatively, if a position is occupied by a few residues with closely related side chains.

§ Peptides known to contain K^d-restricted T-cell epitopes were aligned according to their Tyr residues. Amino-acid residues located outside the nonapeptides predicted to be the naturally processed ones are omitted. Peptides known to be the naturally processed ones are underlined. SYFPEITHI is the prominent self-peptide eluting at fraction 29 in Fig. 1*a*, *b* (see text).

‖ This peptide allows two possibilities for alignments with the motif, taking into account the similar side chains of Ile, Leu, Val and Ala.

¶ The residue at position 10 is in italics. These peptides might be exceptional K^d-restricted epitopes having their C-terminal anchor residues at position 10 instead of 9.

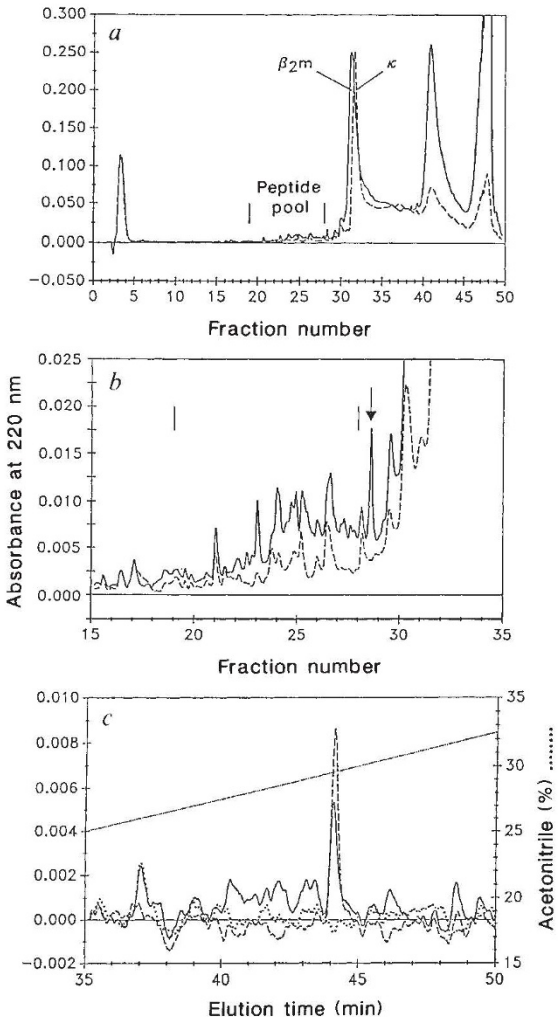

FIG. 1 HPLC separation of immunoprecipitated and TFA-treated K^d molecules. *a*, HPLC profile of TFA-treated material precipitated from P815 lysate with anti-K^d beads (——) or mock-precipitated with irrelevant anti-D^b beads (- - - -). The material eluting beyond fraction 32 is probably MHC light and heavy chains and antibody components. The peak at fraction 48 is detergent. *b*, Enlarged view of the chromatogram in *a* (fraction 15–35). Note several distinct peaks in the K^d precipitate (———), presumably self peptides, which do not appear in the control precipiate (- - - -). The most dominant peak is marked by an arrow (fraction 29) and was determined by sequencing to be SYFPEITHI. *c*, Rechromatography of the dominant self-peptide marked by an arrow in *b*(——), of the corresponding fraction in a control precipitate produced with glycine-coupled beads (· · · ·), and of 500 ng synthetic SYFPEITHI (- - - -).

METHODS. $10–20 \times 10^9$ P815 cells were pelleted and stirred for 30 min with 250 ml 0.5% Nonidet P-40 in phosphate buffered saline (PBS) containing 0.1 mM PMSF at 4 °C. The supernatant (after centrifugation at 250*g*, 5 min, followed by 150,000*g*, 30 min at 4 °C) was passed through a chromatography column (bed volume, 1 ml) filled with glycine-coupled beads, then through a similar column filled with anti-K^d beads, then for a mock-precipitate over anti-D^b beads. Beads removed from all three columns were swirled with 0.1% TFA for 15 min. Supernatants were dried by vacuum centrifugation and separated by reversed-phase HPLC using a Superpac Pep S Column (C2/C18; 5 μm particles, 4.0×250 mm; Pharmacia LKB) and Pharmacia LKB equipment[4]. Eluents: solution A, 0.1% TFA in H_2O (v/v); solution B, 0.1% TFA in acetonitrile. Gradient used in *a* and *b*: 0–5 min, 100% A; 5–40 min, linear increase to 60% B; 40–45 min, 60% B, 45–50 min, decrease to 0% B; flow rate 1 ml min^{-1}, fraction size 1 ml. Fraction size in *c*, 0.5 ml. Individual fractions were collected and dried by vacuum centrifugation. Antibody-coupled and glycine-coupled beads were prepared using cyanogen bromide-activated Sepharose 4B (Pharmacia LKB) according to manufacturer's protocol. K^d-specific antibodies (5 mg each) 20-8-4S (IgG2a, κ; ref. 25) or D^b specific B22-249 (IgG2a, κ; ref. 26) were coupled to 1 ml of beads. β2m, β_2-microglobulin.

sequencing. Identity of this natural self-peptide was confirmed by its coelution with synthetic SYFPEITHI on HPLC (Fig. 1*c*). The sequence fits well with the motif from the pool of fractions 20–28 (Fig. 1*a*, *b*), thus confirming the K^d-restricted peptide motif (Table 1*d*). It will be of interest to see whether alloreactive T cells can recognize this prominent self-peptide, which we estimate, on the basis of comparison with the peak area of the coeluting synthetic peptide, to occupy about 5% of K^d molecules of P815 cells (10,000 SYFPEITHI molecules per cell).

Elution of peptides from K^b and D^b

Detergent lysates from EL4 tumour cells ($H\text{-}2^b$) were immunoprecipitated with K^b-specific and D^b-specific antibodies. Peptides dissociated from MHC molecules were separated by reversed-phase HPLC. Both K^b- and D^b-derived material eluted with profiles roughly similar to the K^d-derived material, with marked differences, however, in the heterogeneous material eluting between fractions 20 and 28 (not shown).

D^b-restricted peptide motif

Pooled fractions 20–28 from the D^b preparation were sequenced (Table 2*a*, *b*). Each of positions 2 to 4 contained several residues. By contrast, cycle 5 gave a strong signal for Asn, which was much weaker for cycle 4 (from 4.2 to 271.4 pmol in the first and from 6.7 to 154.7 pmol in the second experiment). Thus, the dominant residue at position 5 of D^b-eluted self-peptides is Asn. The weak signal for Asp is caused by hydrolysis of Asn to Asp under sequencing conditions. Positions 6–8 contained 5–14 different detectable residues. Position 9 contained a strong signal for Met, an intermediate one for Ile, and a weak one for Leu (all hydrophobic). (The importance of Met or Ile in a D^b-restricted epitope has been reported earlier[17].) Position 10 had no signal, indicating D^b-presented self-peptides to be nonapeptides, as is the D^b-restricted influenza peptide[4]. The consensus motif indicated by these data is shown in Table 2*c*. Comparing this motif with the natural D^b-restricted peptide and with other peptides containing D^b-restricted epitopes shows that Asn at position 5 may be an invariant anchor residue (for definition, see Table 1) of the D^b-restricted peptide motif, not unlike Tyr at position 2 for the K^d-restricted motif. The other residues of the D^b-restricted epitopes vary considerably, with the exception of the Met, and also either Ile or Leu at position 9, which looks like a second anchor position. This anchor is not at position 9 in the (few) known T cell epitopes, except in the naturally processed one.

K^b-restricted peptide motif

Pooled fractions 20–28 from the K^b preparation were sequenced (Table 3*a*, *b*). No strong signal for any residue was at the second position. Position 3 contained a good signal for Tyr, and a weak one for Pro. Position 4 revealed weak signals for five residues. Strong signals for Phe (increase from 1.8 to 50.5 and from 1.5 to 18.3 pmol) and for Tyr make these two residues dominant at position 5. The next two positions contained five or three residue signals. Position 8 had a strong signal for Leu, an intermediate one for Met, and weak ones for Ile and Val. Position 9 showed no increase for any residue, consistent with the length of the known K^b-restricted natural peptide, which is an octamer[5]. Analysis of the consensus K^b restricted motif and comparison with epitopes indicates two anchor positions: Tyr or Phe (both with similar aromatic side chains) at position 5 and Leu, Met, Ile or Val (all with similar hydrophobic side chains) at position 8.

HLA-A2.1 restricted peptide motif

Detergent lysate of human JY cells (HLA-A2.1) was immunoprecipitated with A2-specific antibodies. Peptides dissociated from A2 molecules were separated by HPLC. Fractions 20–28 were pooled and sequenced as for the mouse material (Table 4). The second position contained a strong signal for

TABLE 2 Sequencing of the self-peptide mixture eluted from D^b molecules*

(*a*) Experiment 1

	Amino-acid residues (in pmol)																	
	A	R	N	D	E	Q	G	H	I	L	K	M	F	P	S	T	Y	V
Cycle	Ala	Arg	Asn	Asp	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Tyr	Val
1	257.2	18.2	21.6	1.3	8.1	16.3	99.1	2.3	22.0	21.2	20.3	7.2	33.0	27.5	124.6	43.9	26.9	70.1
2	202.1	7.2	5.4	6.8	7.4	24.7	116.2	0.9	5.4	9.9	6.5	154.1	4.3	8.2	52.7	15.0	5.5	16.0
3	29.9	5.9	5.3	0.0	3.8	5.5	185.1	1.1	106.3	65.8	0.0	8.3	3.8	88.1	8.3	4.7	5.2	73.2
4	18.3	8.1	4.2	4.6	32.4	21.8	49.3	0.8	32.7	21.5	12.4	3.6	2.3	28.8	9.9	18.6	5.0	165.2
5	6.8	2.1	271.4	26.0	8.2	4.3	43.0	0.6	4.7	6.2	2.5	1.3	0.9	11.7	4.5	5.0	1.7	7.6
6	42.1	5.9	29.6	7.1	8.4	7.8	32.6	1.3	18.0	148.4	8.8	1.9	11.3	22.5	7.8	11.8	4.1	23.6
7	21.5	23.4	18.2	24.5	30.4	13.7	22.0	0.7	9.9	16.2	2.4	2.1	3.6	16.4	6.7	54.3	5.1	35.0
8	14.6	10.1	11.3	9.8	23.2	10.3	18.2	0.3	3.0	10.1	4.4	1.3	5.0	9.5	26.5	24.9	12.5	20.7
9	7.5	3.2	7.9	3.2	3.1	1.6	11.2	0.5	8.5	13.7	0.5	7.7	3.0	2.5	2.0	3.3	3.6	3.5
10	2.6	1.1	2.5	2.4	1.9	1.2	12.5	0.3	4.2	8.5	0.4	2.7	1.8	2.1	1.6	1.7	1.9	1.3

(*b*) Experiment 2

1	413.4	45.8	29.7	15.9	14.5	19.6	132.4	4.7	41.5	40.8	48.9	17.2	50.8	26.1	307.7	94.0	47.4	110.1
2	227.4	14.4	7.6	9.3	11.1	25.2	133.8	2.1	8.2	14.5	13.3	169.9	5.6	4.9	71.0	21.6	11.3	22.6
3	39.6	3.3	6.0	6.3	6.0	5.3	172.2	1.2	89.5	56.0	1.6	14.7	4.5	75.4	12.1	5.0	7.6	79.2
4	29.3	16.6	6.7	10.6	34.8	23.0	57.3	0.8	36.3	21.7	17.0	8.1	4.2	33.5	12.5	23.9	7.4	198.9
5	19.9	5.3	154.7	22.2	8.7	4.1	31.1	0.9	4.6	7.0	4.3	2.4	1.7	11.8	5.3	5.0	2.0	13.8
6	42.3	8.4	30.8	15.7	14.6	8.3	28.7	2.3	18.6	124.1	8.2	5.3	11.2	22.1	7.9	10.7	5.6	29.2
7	22.0	24.5	15.4	33.5	29.2	10.5	17.7	1.6	11.3	14.8	3.3	3.7	3.6	14.3	7.5	47.3	6.9	35.5
8	15.8	10.9	10.2	20.9	25.6	8.0	12.6	3.2	3.3	13.6	4.3	2.8	5.1	8.7	20.8	19.3	12.9	23.6
9	8.7	4.3	6.1	13.0	12.1	2.6	8.7	0.3	19.8	26.2	1.2	30.8	3.9	4.4	4.8	5.6	7.2	9.2
10	5.4	3.1	3.9	12.2	8.1	2.0	8.2	0.0	10.1	13.9	0.7	11.6	3.2	3.4	3.0	3.0	7.3	5.9

(*c*) The D^b-restricted peptide motif

	Position											
	1	2	3	4	**5**	6	7	8	**9**†			
Dominant anchor residues					**N**				**M**			
Strong		**M**	**I** **L** **P** **V**	**K** **E** **Q** **V**		**L** **F**			**I**			
Weak	A N I F P S T V	A Q D	G	D T		A Y T V M E Q H I K P S	D E Q V T Y	F H K S Y	L			

Known epitopes, aligned‡												Protein source	Ref.
	A	S	N	E	N	M	E	T	M			Influenza NP366-374	4, 2
	S	G	P	S	N	T	P	P	E	*I*§		Adenovirus E1A	38
	S	G	V	E	N	P	G	G	Y	*C*	*L*§	Lymphocyte choriomeningitis virus GP 272-293	39
	S	A	I	N	N	Y	.	.				Simian virus 40 T 193-211‖	40

* Lysates of EL4 cells were prepared as indicated for the P815 lysate in Fig. 1, and then passed through chromatography columns containing glycine beads, anti-K^b (K9-178, IgG2a, κ; ref. 27) or anti-D^b beads. Anti-D^b beads were treated with 0.1% TFA, and the supernatant was separated by HPLC as for the K^d precipitate in Fig. 1. Fractions 20–28 were sequenced and analysed as in Table 1.

† Anchor positions indicated in bold types.

‡ Alignment was on the Asn residue at position 5. The epitope known to be naturally processed is underlined. For the other peptides, only the parts likely to be the naturally processed epitopes are indicated, although most of the peptides were described as longer ones in the original references.

§ Residues at position 10 and 11 are in italics. These peptides might be exceptional D^b-restricted epitopes having their C-terminal anchor residue at position 10 or 11 instead of 9. Note that both epitopes contain four 'small' amino-acid residues, Gly and Pro, which may influence the linear distance between N and C termini.

‖ With the alignment proposed, positions 7–9 are not covered by the peptide published.

Leu and an intermediate one for Met. Positions 3–5 had 6–8 residues each. Position 6 contained Val, Leu, Ile, and Thr. Each of the following two positions had three signals. Position 9 had a strong Val and a weak Leu signal. Position 10 showed no increase for any residue, indicating A2-restricted epitopes to be nonapeptides. Anchors appear to be Leu or Met at position 2 and Val or Leu at position 9. Some of the peptides reported to contain A2-restricted epitopes can be aligned to the motif, whereas others can be aligned only partially (Table 4*c*). The existence of variant A2 molecules typing as A2 by serology may cause the poor alignment to the motif of some of the peptides. The epitope content of some of these peptides has also not been formally established.

Discussion

Sequencing of the self-peptide mixtures from the four MHC class I molecules H-$2K^d$, H-$2K^b$, H-$2D^b$ and HLA-A2 indicated a distinct allele-specific peptide motif presented by each class I molecule. K^d-, D^b- and A2-presented peptides are nonamers, whereas K^b-presented peptides seem to be octamers; the corresponding peptide motifs contain two anchor positions occupied by a fixed residue or by one of a few residues with closely related

TABLE 3 Sequencing of the self-peptide mixture eluted from K^b molecules*

(*a*) Experiment 1

	Amino-acid residues (in pmol)																	
	A	R	N	D	E	Q	G	H	I	L	K	M	F	P	S	T	Y	V
Cycle	Ala	Arg	Asn	Asp	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Tyr	Val
1	978.7	26.3	49.2	55.8	39.0	23.1	514.9	20.9	167.5	167.2	189.8	50.3	116.7	118.2	120.8	365.2	136.0	352.5
2	345.5	3.9	37.3	41.8	23.5	20.3	475.2	8.9	44.5	43.1	72.6	12.6	25.4	51.0	253.1	80.5	50.1	93.5
3	129.0	1.4	14.7	37.0	17.7	9.8	358.8	5.9	8.2	19.0	26.9	4.1	6.0	32.5	56.2	20.0	75.6	25.9
4	52.1	3.5	10.8	45.3	38.0	9.2	246.7	5.0	4.9	7.0	17.7	2.4	1.8	14.6	23.0	13.4	12.0	16.4
5	18.9	1.3	5.5	34.7	12.0	3.6	128.2	2.8	1.9	4.7	3.8	1.6	50.5	6.7	8.9	4.6	33.2	4.9
6	16.2	0.8	5.6	32.7	13.0	3.7	77.9	2.4	3.1	3.5	3.9	0.9	4.5	7.3	9.2	18.3	7.3	6.2
7	9.9	0.9	14.9	30.4	9.5	6.6	51.3	0.6	0.0	3.4	9.2	0.5	1.9	4.7	6.1	10.7	3.5	3.4
8	6.0	1.4	5.1	22.7	6.0	3.3	29.2	0.8	1.4	13.5	1.8	2.1	1.0	3.8	4.1	3.1	2.5	3.6
9	4.6	1.5	2.6	19.9	4.5	2.3	21.1	0.9	0.9	6.9	1.0	1.5	1.0	3.0	3.7	2.2	1.9	2.1
10	3.9	0.5	1.9	17.5	3.7	2.1	17.5	1.0	0.5	4.0	0.8	0.9	1.2	2.8	3.5	1.8	2.0	1.5

(*b*) Experiment 2

1	42.4	1.1	5.2	3.0	7.8	17.1	44.6	0.3	11.3	12.6	12.1	3.8	6.2	7.6	44.2	18.1	6.8	26.2
2	24.0	0.2	9.4	2.8	5.1	8.0	42.5	0.5	4.7	6.3	4.0	1.3	3.7	3.5	14.9	10.3	3.1	6.9
3	10.4	0.3	2.1	2.6	3.9	4.0	25.1	0.7	2.8	7.9	2.1	0.9	3.6	9.8	3.0	3.3	16.7	10.0
4	9.6	1.3	2.7	5.7	7.5	4.1	24.5	0.2	1.5	5.0	6.3	0.7	1.5	5.9	3.0	5.9	2.7	4.5
5	5.8	0.8	1.8	2.8	3.3	2.5	14.2	0.5	0.2	3.9	1.7	0.4	18.3	3.5	1.3	2.0	20.8	2.2
6	8.6	0.2	2.3	2.7	6.3	2.7	9.2	0.0	1.0	2.4	1.5	0.4	2.3	3.2	2.7	5.2	3.6	2.4
7	5.0	0.1	8.2	3.3	3.9	4.2	10.4	0.6	0.4	2.3	7.2	0.1	1.2	2.1	1.9	2.8	1.9	1.2
8	4.0	0.1	3.1	2.0	2.6	1.7	6.9	0.2	0.2	13.8	1.6	1.0	0.8	1.1	0.7	1.3	1.1	2.2
9	4.5	0.1	1.1	2.1	3.6	1.9	5.9	1.4	0.0	7.7	0.9	1.0	0.9	1.3	0.3	1.3	0.8	1.7
10	3.9	1.7	0.3	4.5	3.0	1.4	5.4	0.2	0.0	3.9	0.6	0.6	0.6	1.1	0.6	1.1	0.8	1.1

(*c*) The K^b-restricted peptide motif

	Position							
	1	2	3	4	**5**	6	7	**8**†
Dominant anchor residues					**F** **Y**			**L**
Strong			**Y**					**M**
Weak	R I L S A	N	P	R D E K T		T I E S	N Q K	I V

Known epitopes, aligned‡									Protein source	Ref.
	R	G	Y	V	Y	Q	G	L	Vesicular stomatitis virus NP 52-59	5
	S	I	I	N	F	E	K	L	Ovalbumin 258–276§	41
	A	P	G	N	Y	P	A	L	Sendai Virus NP 321-332	42

* The anti-K^b beads loaded with EL4 lysates described in Table 2 were treated with 0.1% TFA as were the anti-K^d beads in Fig. 1. Extracted material was HPLC-separated as before, and fractions 20–28 were pooled, sequenced, and analysed as in Table 1.

† Anchor positions are in bold types.

‡ Alignment was on the Leu residue at position 8. The only known naturally processed epitope is underlined. For the others, only the postulated natural epitopes are indicated; additional residues contained in the original references have been omitted.

§ The published ovalbumin peptide did not contain the Ser residue at position 1.

side chains. These anchor positions are not at the same place in the different motifs; they are at position 5 and 9 (D^b), or 2 and 9 (K^d, A2), or 5 and 8 (K^b). The C-terminal anchor residues of all motifs are hydrophobic. For H-$2L^d$ (data not shown), one anchor residue was Pro at position 2. (The entire motif could not be sequenced, as not enough material eluted from L^d molecules.) The residues not at anchor positions can be fairly variable; some, however, seem to be preferentially occupied; for example, Pro is prominent at position 4 of the K^d motif, Tyr at position 3 of the K^b motif, and hydrophobic residues predominate at positions 3 of the D^b motif and 6 of the A2 motif. Because of technical limitations of our unconventional approach, less abundant amino-acid residues may have escaped detection; Cys and Trp residues were not detected at all. Comparison of the motifs with known epitopes indicates that some residues have been missed by sequencing the self-peptide mixtures. A minor population of peptides not adhering to the motif-specific lengths also cannot be excluded. The K^d- and D^b-restricted epitopes whose positions 9 do not fit with the hydrophobic anchor of the respective motifs would have residues at position 10 or 11 that would fit (Tables 1*d*, 2*c*).

Our observations match well the structure of the peptide-binding cleft of MHC class I molecules[3,6]. With HLA-A2, the cleft has pockets of a size that would accommodate specific amino-acid side chains, for example Leu[3,6]. Leu dominantly occupies position 2 of the A2 motif. Another pocket should exist in A2 molecules to accommodate the side chains of Val and Leu at position 9 of A2-restricted epitopes. Co-crystallizing material not from the A2 sequence and bound to the cleft showed extensions (possibly Leu and Val side chains) fitting the A2 pockets[3,6]. With the Aw68 crystal, which differs from A2 at 13 amino-acid residues, pockets of different location and of different shape from that of A2 were found[6]. Therefore, different MHC class I alleles differ in the location and shape of pockets in the cleft likely to be able specifically to accommodate certain amino-acid side chains.

Thus, the allele-specific pockets in MHC crystals and the side chains of the allele-specific anchor residues we find in the sequenced self-peptides are likely to reflect complementary structures. From this, we deduce that the structure of K^d (mouse MHC molecules have not been crystallized yet) possesses a pocket for the aromatic side chains of Tyr or Phe at one end of the cleft and another one at the opposite end, able to accommodate Ile, Leu, or Val (all hydrophobic with side-chain methyl groups). Similarly, K^b should have a pocket close to the cleft's centre for the aromatic side chains of Tyr or Phe, and another

TABLE 4 Sequencing of the self-peptide mixture eluted from A2.1 molecules*

(*a*) Experiment 1

	Amino-acid residues (pmol)																	
Cycle	A Ala	R Arg	N Asn	D Asp	E Glu	Q Gln	G Gly	H His	I Ile	L Leu	K Lys	M Met	F Phe	P Pro	S Ser	T Thr	Y Tyr	V Val
1	172.6	0.0	31.9	25.7	44.8	125.9	112.4	2.8	144.4	123.8	60.0	30.7	63.3	117.9	75.9	49.0	50.3	104.9
2	42.5	0.0	16.2	14.1	25.6	53.1	44.7	1.6	69.6	511.0	15.5	71.0	10.5	38.7	16.2	16.1	12.2	86.5
3	99.8	0.0	9.5	18.3	12.3	20.4	31.8	11.1	51.5	110.8	5.8	55.7	19.4	30.4	12.0	8.7	20.9	46.0
4	36.0	0.6	12.7	26.4	59.5	21.7	56.2	1.3	10.4	22.7	24.6	5.2	5.2	52.4	10.9	14.0	5.2	28.8
5	35.1	0.1	13.4	18.6	28.1	19.8	55.6	2.8	21.4	23.9	47.2	4.1	6.2	39.1	7.5	10.5	11.6	29.0
6	30.3	0.9	16.8	14.1	21.4	17.3	28.5	1.4	68.1	43.4	14.7	4.4	5.8	40.8	9.2	20.3	5.0	106.2
7	42.1	0.0	11.7	9.5	27.2	21.8	19.0	3.2	36.3	27.3	7.9	5.7	8.0	54.1	5.4	13.6	14.0	62.8
8	37.9	0.3	13.4	8.1	37.3	24.3	21.1	1.8	11.6	15.1	33.8	3.4	5.1	22.3	8.8	17.9	10.2	22.4
9	23.3	0.0	5.1	6.0	15.7	10.5	14.8	0.7	11.5	27.5	8.7	3.1	2.7	11.9	5.6	6.7	5.1	60.2
10	12.0	0.7	2.6	4.4	6.5	5.2	10.2	0.4	4.5	12.1	4.5	1.0	1.8	7.1	2.7	3.2	2.3	20.4
(*b*) Experiment 2																		
1	110.8	10.8	4.0	3.1	10.0	14.5	55.7	0.2	60.3	44.4	10.8	8.2	37.5	20.3	27.4	14.6	19.8	48.0
2	13.4	1.6	2.0	1.9	6.8	11.0	9.0	0.0	37.9	302.7	0.0	26.2	5.0	6.3	4.4	4.5	3.3	26.5
3	62.4	3.5	5.0	9.1	4.9	10.0	12.6	0.1	35.7	71.5	0.0	24.5	13.8	13.4	8.9	4.8	17.9	19.6
4	16.9	2.2	4.5	8.9	25.3	7.9	24.5	0.1	6.2	10.3	2.8	1.3	2.0	22.1	4.9	5.0	1.8	9.3
5	22.3	1.6	6.8	8.6	14.3	9.9	31.8	0.0	16.6	15.1	8.2	1.9	4.0	16.3	4.5	4.6	5.7	18.3
6	10.6	1.3	6.6	3.6	6.4	6.2	10.1	0.1	38.7	27.1	0.0	1.4	2.7	12.6	3.2	6.1	1.3	39.2
7	19.2	1.0	4.7	2.5	7.2	9.0	5.6	0.2	22.3	16.1	0.0	1.9	3.9	17.4	1.9	3.5	3.6	27.2
8	13.4	1.2	3.1	1.3	7.9	6.3	6.9	0.3	4.7	6.7	3.0	0.6	2.0	5.1	2.2	4.9	1.6	5.3
9	5.7	0.5	0.9	0.8	2.9	2.0	2.7	0.2	3.8	11.5	0.4	0.3	0.6	2.0	1.0	1.1	0.4	10.8
10	2.9	0.6	0.5	0.5	1.0	0.9	1.8	0.3	1.6	5.1	0.4	0.3	0.3	0.8	0.4	0.3	0.2	3.6

(*c*) The HLA-A2-restricted peptide motif

	Position								
	1	**2**	3	4	5	6	7	8	**9**†
Dominant anchor residue		**L**							**V**
Strong		**M**		**E** **K**		**V**		**K**	
Weak	I L F K M Y V		A Y F P M S R	G P D T	I K Y N G F V H	I L T	A Y H	E S	L

Reported epitopes, aligned‡											Protein source	Ref.
	I	L	K	E	P	V	H	G	V		HIV reverse transcriptase 461–485	43
	G	I	L	G	F	V	F	T	L		Influenza matrix protein 57–68	44
	I	L	G	F	V	F	T	L	T	*V*¶	Influenza matrix protein 57–68	44
	F	L	Q	S	R	P	E	P	T		HIV Gag protein 446–460§	46
	A	M	Q	M	L	K	E		.		HIV Gag protein 193–203§	46
	P	I	A	P	G	Q	M	R	E		HIV Gag protein 219–233§	46
	Q	M	K	D	C	T	E	R	Q		HIV Gag protein 418–443§	46

* Detergent lysate of JY cells (HLA-A2.1)[45] prepared as indicated for P815 cells in Fig. 1 was passed through a column containing glycin beads, then through a column containing anti-A2 (BB7.2, IgG2b; ref. 28) beads. Loaded anti-A2 beads were treated with 0.1% TFA, and the material in the supernatant was separated by HPLC as for the K^d preparation in Fig. 1. Fractions 20 to 28 were pooled and sequenced as in Table 1.

† Anchor positions are in bold types.

‡ Alignment was on the anchor at position 2 of the motif. Only the nonapeptides fitting to the motif are indicated, although all of the peptides were described as longer ones in the original references. For the influenza matrix peptide, an alternative alignment is shown, which employs an Ile residue at position 2. Ile and Leu have very similar side chains. With this alignment, position 9 of the peptide fits with the C-terminal anchor of the motif. In addition, position 6 of the peptide then fits well with position 6 of the motif, which may represent an auxiliary anchor of this motif preferentially for hydrophobic residues.

§ All four HIV Gag peptides can only partially be aligned with the motif. These peptides induced low or intermediate lysis of target cells at micromolar concentrations[46], whereas natural epitopes induce lysis at picomolar concentrations[4]. All four peptides were recognized by the same cytotoxic T lymphocyte line[46].

¶ The residue at position 10 is in italics. This peptide might be an exceptional A-2 restricted epitope having its C-terminal anchor residue at position 10 instead of 9.

at the very end of the cleft for Leu or Met, which have hydrophobic side chains similar to each other. For D^b we postulate a central pocket exclusively for Asn. A second pocket for Met, Ile, Leu, or Val should be at one end of the D^b-cleft. This pocket might be less specific, as some peptides reported to contain D^b-restricted epitopes do not fit to the motif at this position. Alternatively, these peptides may represent examples of a minority of epitopes having their C-terminal hydrophobic anchor residue at position 10 or 11 (Table 2*c*). L^d molecules should contain a pocket for Pro at one end of the groove.

The mechanism by which MHC-restricted epitopes are produced is unclear, even when they stem from known proteins. As all epitopes appear to have hydrophobic C termini, whereas the N termini are highly variable, the hydrophobic end might be a site for an MHC-independent protease[8], producing larger precursor peptides, whose N termini are trimmed after binding to MHC[8].

Strictly allele-specific peptide motifs seem to be contradicted by data obtained from assays measuring MHC binding to solid-phase peptides (for example, refs 19, 20); later studies however suggest that this reflects the artificial situation in which these assays are done. We have also shown that artificial peptides (not adhering to MHC restricted motifs) binding to MHC molecules, and even inducing T-cell responses are not processed

and presented by cells expressing the respective protein[21]. From this we conclude that more peptides can bind to MHC molecules than are processed by cells. Taking into account the strong adherence of naturally processed peptides to the MHC-restricted motifs, we also conclude that these motifs are not 'peptide binding motifs', but motifs representing the outcome of processing, which is likely to include MHC-dependent and MHC-independent protease activities[8,9] and transport mechanisms[22–24]. Intracellular binding to MHC molecules of epitopes correctly processed without participation of MHC molecules cannot entirely be excluded, although it is unlikely, given our inability to find intracellularly correctly processed epitopes independent of MHC[8]. Thus peptide binding to MHC molecules is a necessary requirement for a peptide being an MHC-restricted epitope but is not sufficient on its own.

Knowledge of the peptide motifs of individual MHC alleles should help exact T-cell epitope predictions (K.F. *et al.*, manuscript submitted; K. Deres *et al.*, unpublished) and help with synthetic or recombinant vaccine development and potentially also for intervention in autoimmune diseases or graft rejection. □

Received 6 February; accepted 4 April 1991.

1. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248,** 701–702 (1974).
2. Townsend, A. R. *et al. Cell* **44,** 959–968 (1986).
3. Bjorkman, P. J. *et al. Nature* **329,** 512–518 (1987).
4. Rötzschke, O. *et al. Nature* **348,** 252–254 (1990).
5. Van Bleek, G. M. & Nathenson, S. G. *Nature* **348,** 213–216 (1990).
6. Garrett, T. P. J., Saper, M. A., Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *Nature* **343,** 692–696 (1989).
7. Rötzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249,** 283–287 (1990).
8. Falk, K., Rötzschke, O. & Rammensee, H.-G. *Nature* **348,** 248–251 (1990).
9. Shimonkevitz, R., Kappler, J., Marrack, P. & Grey, H. *J. exp. Med.* **158,** 303–316 (1983).
10. Demotz, S., Grey, H. M., Appella, E. & Sette, A. *Nature* **343,** 682–684 (1989).
11. Bjorkman, P. J. *et al. Nature* **329,** 506–512 (1987).
12. DeLisi, C. & Berzofsky, J. A. *Proc. natn. Acad. Sci. U.S.A.* **82,** 7048–7052 (1985).
13. Rothbard, J. B. & Taylor, W. R. *EMBO J.* **7,** 93–100 (1988).
14. Cornette, J. L., Margalit, H., DeLisi, C. & Berzofsky, J. A. *Meth. Enzym.* **178,** 611–633 (1989).
15. Sette, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **86,** 3296–3300 (1989).
16. Maryanski, J. L., Verdini, A. S., Weber, P. C., Salemme, F. R. & Corradin, G. *Cell* **60,** 63–72 (1990).
17. Bastin, J., Rothbard, J., Davey, J., Jones, I & Townsend, A. *J. exp. Med.* **165,** 1508–1523 (1987).
18. Bjorkman, P. J. & Davis, M. M. *Cold Spring Harb. Symp. quant. Biol.* **54,** 365–374 (1989).
19. Bouillot, M. *et al. Nature* **339,** 473–475 (1989).
20. Frelinger, J. A., Gotch, F. M. Zweerink, H., Wain, E. & McMichael, A. J. *J. exp. Med.* **172,** 827–834 (1990).
21. Schild, H., Rötzschke, O., Kalbacher, H. & Rammensee, H.-G. *Science* **247,** 1587–1589 (1990).
22. Townsend, A. *et al. Nature* **340,** 443–448 (1989).
23. Elliott, T., Townsend, A. & Cerundolo, V. *Nature* **348,** 195–197 (1990).
24. Cerundolo, V. *et al. Nature* **345,** 449–452 (1990).
25. Rüsch, E., Kuon, W. & Hämmerling, G. J. *Trans. Proc.* **15,** 2093–2096 (1983).
26. Lemke, H., Hämmerling, G. J. & Hämmerling, U. *Immunol. Rev.* **47,** 175–206 (1979).
27. Ozato, K. & Sachs, D. H. *J. Immun.* **126,** 317–321 (1981).
28. Parham, P. & Brodsky, F. M. *Hum. Immun.* **3,** 277–299 (1981).
29. Taylor, P. M., Davey, J., Howland, K., Rothbard, J. B. & Askonas, B. A. *Immunogenetics* **26,** 267–272 (1987).
30. Braciale, T. J. *et al. J. exp. Med.* **166,** 678–692 (1987).
31. Braciale, T. J., Sweetser, M. T., Morrison, L. A., Kittlesen, D. J. & Braciale, V. L. *Proc. natn. Acad. Sci. U.S.A.* **86,** 277–281 (1989).
32. Kuwano, K., Braciale, T. J. & Ennis, F. A. *FASEB J.* **2,** 2221 (1988).
33. Maryanski, J. L., Pala, P., Cerottini, J. C. & Corradin, G. *J. exp. Med.* **167,** 1391–1405 (1988).
34. Maryanski, J. L., Pala, P., Corradin, G., Jordan, B. R. & Cerottini, J. C. *Nature* **324,** 578–579 (1986).
35. Sibille, C. *et al. J. exp. Med.* **172,** 35–45 (1990).
36. Romero, P. *et al. Nature* **341,** 323–326 (1989).
37. Weiss, W. R. *et al. J. exp. Med.* **171,** 763–773 (1990).
38. Kast, W. M. *et al. Cell* **59,** 603–614 (1989).
39. Oldstone, M. B. A., Whitton, J. L., Lewicki, H. & Tishon, A. *J. exp. Med.* **168,** 559–570 (1988).
40. Tevethia, S. S. *et al. J. Virol.* **64,** 1192–1200 (1990).
41. Carbone, F. R. & Bevan, M. J. *J. exp. Med.* **169,** 603–612 (1989).
42. Schumacher, T. N. M. *et al. Cell* **62,** 563–567 (1990).
43. Walker, B. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **86,** 9514–9518 (1989).
44. Gotch, F., McMichael, A. & Rothbard, J. *J. exp. Med.* **168,** 2045–2057 (1988).
45. Santos-Aguado, J., Crimmins, M. A. V., Mentzer, S. J., Burakoff, S. J. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **86,** 8936–8940 (1989).
46. Claverie, J. M. *et al. Eur. J. Immun.* **18,** 1547–1553 (1988).
47. Falk, K. *et al. J. exp. Med.* (in the press).

ACKNOWLEDGEMENTS. We thank J. Klein for support, A. Townsend for cell lines, S. Faath for technical assistance, and L. Yakes for typing the manuscript. This work was supported by Sonderforschungsbereich 120 and 323.

LETTERS TO NATURE

Statistical evidence for a galactic origin of gamma-ray bursts

J. L. Atteia*, C. Barat*, E. Jourdain*, M. Niel*, G. Vedrenne*, N. Blinov*, A. Chernenko†, V. Dolidze†, A. Kozlenkov†, A. Kuznetsov†, I. G. Mitrofanov†, A. Pozanenko†, R. Sunyaev† & O. Terekhov†

* Centre d'Etude Spatiale des Rayonnements, 9 Avenue du Colonel Roche, BP 4346, 31029 Toulouse Cedex, France

† Space Research Institute, Profsoyuznaya 84/32, 117810 Moscow, USSR

ASTRONOMICAL bursts of gamma-rays (GRBs) were first discovered 20 years ago, and ~100 are recorded every year by satellite-borne instruments. Bursts last for at most a few seconds, recur only on a timescale of years, if at all, and come from objects which have remained undetected at all other wavelengths. It has been impossible to establish a distance scale for GRBs, and no association with known astronomical objects has been demonstrated. Here we analyse the spatial distribution of GRBs[1,2] with a view to understanding their true radial distribution. Our data consist of three samples, totalling 244 GRBs, obtained by three French–Soviet experiments flown on the Venera 13 and 14 and the Phobos missions. We conclude that the underlying GRB distribution is not uniformly distributed in space, but falls off with distance. Our analysis of the weak sources in particular suggests that GRBs are associated with the galactic plane.

There has recently been interest in using the V/V_{max} distribution to study the radial distribution of the sources[3,4], where for a given burst V is the volume of the smallest sphere containing the source, and V_{max} is the maximum volume accessible to the instrument. This ratio is defined as follows: if D is the (unknown) distance of the GRB source and D_{max} the maximum distance at which it can be detected, then $V/V_{max} = (D/D_{max})^3$, and as the number of detected counts, C, scales as D^{-2}, we can write $V/V_{max} = (C_{max}/C_{min})^{-3/2}$ where C_{max} is the maximum number of counts and C_{min} the minimum number required to detect a burst. The V/V_{max} distribution is free of instrumental selection effects as long as C_{max} and C_{min} are calculated in the same time intervals and energy range as those used to trigger the experiment[3] (detailed analysis shows that some bias can still affect the V/V_{max} distribution; Hartmann *et al.*, manuscript in preparation). If the number of sources per unit volume is constant, corresponding to an infinite uniform density population, the V/V_{max} values are uniformly distributed between 0 and 1. Here we present the V/V_{max} distribution for bursts detected by the Signe experiments on the Soviet Venera 13 and 14 probes, and by the Lilas and Apex experiments on the Phobos probe. We use both the mean value $\langle V/V_{max}\rangle$ and a Kolmogorov–Smirnov (KS) test to compare the observed V/V_{max} distribution with a uniform one. Two of the three data sets exhibit a significant deviation from a uniform distribution, which may be interpreted as a deficit of sources at large distances. We show that the faintest sources (in the V/V_{max} sense) seem to be concentrated in the galactic plane, evidence which supports a galactic-disk distribution of GRB sources. The V/V_{max} analysis has been used previously to study the radial distribution of burst samples[4–7], but here we associate a low $\langle V/V_{max}\rangle$ value with the observation of angular anisotropy.

Table 1 gives the main characteristics of the three experiments used here (details available for Signe[8] and in preparation for Lilas and Apex (C. Barat *et al.*, manuscript in preparation)). The Signe experiment had four identical detectors, the Lilas