Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 7, No. 1, pp. 19-24, 1997

0960-894X/97 $17.00 + 0.00

PII: S0960-894X(96)00579-3

SAR FOR MHC CLASS II BINDING TETRAPEPTIDES: CORRELATION WITH POTENTIAL BINDING SITE

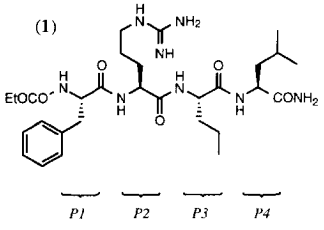
Barry R. Cunningham,[a] Meheryar Rivetna,[a] Richard L. Tolman,[a] Stacey J. Flattery,[b] Elizabeth A. Nichols,[b] Cheryl D. Schwartz,[b] Linda S. Wicker,[b] Jeffrey D. Hermes[b] and A. Brian Jones.[a,1,*]

Departments of Medicinal Chemistry[a] and Immunology & Inflammation[b]
Merck Research Laboratories, R50G-146, P.O.Box 2000, Rahway, NJ, 07065.

Abstract: Comparison of SAR for a tetrapeptide inhibitor of MHC class II with existing information endorses a binding mode consistent with naturally loaded full length peptides.

The acquisition of antigenic peptides by MHC proteins is a critical step in the development of a directed immune response. The pathway involving MHC class II proteins provides the mechanism for selecting and presenting T-cell epitopes from exogenous foreign protein.[2] Although the level of understanding concerning this process has improved dramatically over the last 5 years, the recent exposé of a new and important protein player serves to emphasize the still developing nature of this understanding.[3] In the terminal stage of presentation, ~13-mer peptides complexed to class II protein are exposed on the surface of professional antigen presenting cells for recognition by cognate T-cells through the T-cell receptor. An important milestone in understanding this step was the determination of the co-crystal structure of a single allele of class II (designated DR(B1*0101) or DR1) with a single antigenic peptide (haemaglutinin peptide $HA_{307-319}$, PKYVKQNTLKLAT).[4] Aberrant presentation of autoantigen has been implicated in the induction and maintainance of autoimmune disease[5] and suggests potential for therapeutic intervention with selective inhibitors of class II ligation.[8] We have recently shown that tetrapeptide (**1**)[6] inhibits the binding of full length peptide to purified DR1with nanomolar IC_{50},[7] raising the prospect of allele selective inhibition of epitope loading. On the basis of comparisons of this tetrapeptide with $HA_{307-319}$ and known DR1 binding motifs we made the assumption that (**1**) binds in a manner analogous to 'full length' peptides and that it does so in the N-terminal sector of the class II binding groove corresponding to residues 309-312 (YVKQ) of $HA_{307-319}$. In the course of evaluating the tetrapeptide we have generated some support for this assumption and it this we wish to describe.

(1)



C-terminal truncation of the lead tetrapeptide resulted in loss of binding affinity (Table 1, pairs **3/4**, **5/6**, **8/9**). This correlates with the proposed role of HA_{312} (P4 of tetrapeptide) as one of 4 minor contact residues of the full length peptide,[4] one which might play a more significant role in the shortened inhibitor. Were the tetrapeptide located as suggested, the N-terminal cap would lie outside the binding groove and hence make no binding contribution with the exception of an immediately amino proximal H-bond acceptor that would contact $His\beta_{81}$ replacing the C=O of HA_{308}.[4] Indeed, variation of the N-terminal group was of rather limited consequence (Table 2), with the exception of an apparent limitation on gross bulk (**18**).

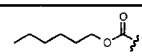
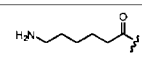
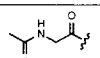
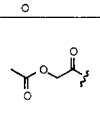
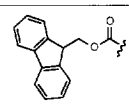
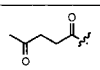
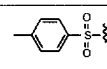
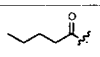
Table 1: IC_{50}s of EtOCO-$(Xaa)_n$-NH_2 Peptides* against DR1/nM^

Compd	Sequence	IC_{50} 20′	IC_{50} 5 H	Compd	Sequence	IC_{50} 20′	IC_{50} 5 H
1	FRNvaL	1.4	17	5	ChaRNvaL†	1.0	2.5
2	(*D*-F)RNvaL	1,500	28,500	6	ChaRNva	2,100	11,500
				7	(*D*-Cha)RNvaL	2,000	30,000
3	FVAL	28	420	8	ChaVAL	1.4	9.9
4	FVA	>50,000	>50,000	9	ChaVA	45,000	>50,000

*In all tables, amino acids are (*L*)-amino acids unless indicated. †20′/5 H IC_{50}s against DR4 are 150 nM/2,500 nM
^IC_{50}s are an average of 2 or more determinations. Generally, we did not consider differences of less than 4-fold to be significant.

In the context of DR1 (and highly homologous alleles, e.g., DR(B1*0401) or DR4) an amino terminal hydrophobic residue (typically F, Y or W) is known to be the dominant feature of the binding motif.[9] Tyr_{309} of $HA_{307\text{-}319}$, for example, provides a disproportionate contribution to overall binding affinity.[10] One might therefore expect variation of the P1 residue of the tetrapeptides to have a notable impact.[11] A selection of P1 side chain variants are presented in Table 3.[12] Ala substitution indicated that hydrophobicity is indeed required for the tetrapeptides (**36** vs **1**).[6] Compounds **21-25** plumb the depth of the P1 interaction. It is clear that natural Phe is optimal is in this regard, correlating with the proportions of the site apparent from the crystal structure. Groups affecting lateral bulk (**33**, **38-40**) were in some cases very well tolerated (e.g., **33** vs **3**, **39** vs **1**). One would not necessarily have expected this from a crystallographic standpoint, since Tyr_{309} appears to fit its pocket quite snugly. On the other hand epitopes bearing Trp in this position are well known.[10] Although this could indicate

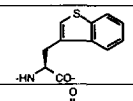
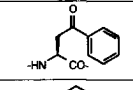
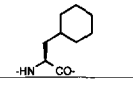
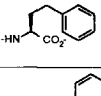
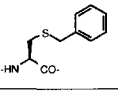
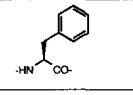
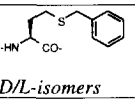
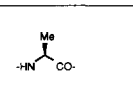
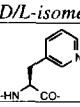
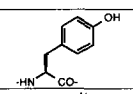
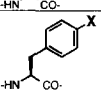
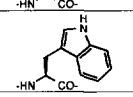
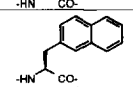
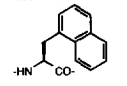
Table 2: IC_{50}s of R-FVAL-NH_2 Peptides against DR1/nM

Compd	R	IC_{50} 20′	IC_{50} 5 H	Compd	R	IC_{50} 20′	IC_{50} 5 H
3		28	420	15		32	17,000
10		22	290	16		170	1600
11		81	-	17		43	500
12		46	390	18		>50,000	>50,000
13		25	430	19		200	1000
14		140	10,000	20	H	1800 (45 min)	-

an alternate binding mode, it more likely reflects a degree of lateral flexibility in the class II protein, a feature perhaps logically suited to its role. Restricted homo-Phe (**34**) was not well accepted, indicating either limits to this flexibility or a distaste for mildly polar functionality. In HA, this anchor residue is Tyr.[4] Various 4-substituted phenylalanines (**27-32**), and Tyr itself (**37**), suggest that a certain leniency to substitution is indeed tolerated at this

position, though not providing marked benefits in any case. Utilizing the hydrophobicity increase of the Phe→Cha transition is well rewarded (**35** vs **3**), the flexibility of the site being adequate to accomodate the added bulk. As befits a hydrophobic site, a basic residue is not welcome (**26**). *R*-configuration residues at this position were not tolerated (Table 1, **2/7**).

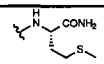
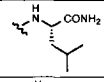
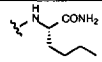
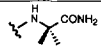
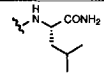
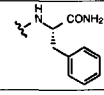
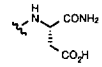
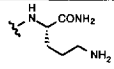
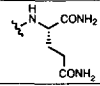
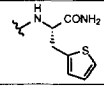
Table 3: IC_{50}s of EtOCO-P1-$(Xaa)_3$-NH_2 Peptides against DR1/nM

Compd	P1	IC_{50} 20′	IC_{50} 5 H	Compd	P1	IC_{50} 20′	IC_{50} 5 H
	P1-VAL			33		23	85
21		65,000	300,000	34		5,700	45,000
3		28	420	35		1.4	9.9
22		160	1,700		**P1-RNvaL**		
23		270	7,000	1		1.4	17
24 25	 *D/L-isomers*	2,100 12,000	>50,000 >50,000	36		8,900	38,000
26		5,300	28,000	37		6.5	63
				38		31	320
27 28 29 30 31 32	X = OMe X = OEt X = OAc X = Cl X = NH_2 X = NO_2	46 17 240 19 540 140	180 75 1,900 73 3,300 600	39 40	 	5.3 16	35 140

As mentioned above, in this frame the P4 residue would correspond to a minor contact residue, Gln_{312} of HA. At this site Gln is clearly not optimal (**45/52**) and the Leu of our starting tetrapeptide,[6] while not unique (**49** vs **41**), does a good job. Inspection of the crystal structure allows rationalization of much of the negative data for P4 substitution (for example, the apparent steric blockage of α-disubstitution (**42/43**) and hydrophobic bulk (**44/50**)[13]) and it is consistent with the view of the site as a rather shallow linear valley. The viability of cyclo-Lys (**53**), interesting from the point of view of firming carboxypeptidase stability, also appears structurally reasonable. That a cationic residue is wholly deleterious (**46**) correlates with known SAR from full length peptides and is related to the proximity of the $Arg\beta_{71}$ sidechain.[14] The influence of the P4 residue on allele selectivity also

provides compelling positional evidence. An anionic residue at P4 favours DR4 at the expense of DR1 (e.g., **51** vs **3**) consistent with its proximity to polymorphic residue β13 (DR1 Phe, DR4 His).[6,15] DR1 and DR4 are otherwise homologous in the P1-P4 region and high affinity DR4 inhibitors can be extrapolated from the DR1 ligands (e.g., 20′/5H IC_{50}s for EtOCO-ChaRNvaD-NH_2 are 2,000 nM/21,000 nM against DR1 and 45 nM/520 nM against DR4; compare to **5**).

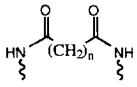
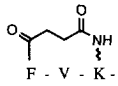
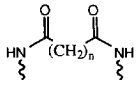
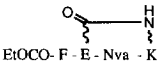
Table 4: IC_{50}s of EtOCO-F-P2-A-P4 Peptides against DR1/nM

Compd	P4	IC_{50} 20′	IC_{50} 5 H	Compd	P4	IC_{50} 20′	IC_{50} 5 H
	P2 = Lys			48		230	910
41		67	1,100	49		81	600
42		27,000	>50,000		**P2 = Val**		
43		730	1,600	3		28 8,200†	420
44		920	5,000	50		920	50,000
45		15,000	43,000	51		>50,000 1,200†	>50,000 50,000†
46		48,000	50,000	52		2,700	26,000
47		1,500	12,000	53		12	82

†IC_{50} against DR4

One of the primary features of the peptide/class II complex is the way the peptide is drawn out in an extended conformation (type II polyproline helix).[4] Denying access to this configuration ought to abrogate binding of the tetrapeptide if it is indeed a 'groove binder'. We assessed the impact of cyclizing tetra- and pentapeptides in increasingly tighter ring structures that would be expected to progressively disrupt linearity (Table 5). A corresponding rough trend in IC_{50} was indeed observed (**56-60**). Cyclic peptide **56** is not impaired relative to the parent peptide (**54**),[16] presumably indicating that the 28-membered ring is sufficient to span the 14Å εLys-εLys distance between these termini observed in the co-crystal without perturbing backbone conformation. At the other end of the scale the IC_{50} for peptide **60** (17-membered ring) is approximately 2 orders of magnitude higher than the parent. Finally, two other cyclization modes were examined (**61/62**). The high IC_{50}s are consistent with enforcing an undesirable twist in the backbone.

Table 5: IC_{50}s of Cyclized Peptides (C-terminal amides) against DR1/nM

Compd	IC_{50} 20′	IC_{50} 5 H	Compd	IC_{50} 20′	IC_{50} 5 H
54: EtOCO-KFVKL	7.1	59	55: EtOCO-FVKL	32	190
 EtOCO-K -- F -- V -- K - L			 F - V - K - L		
56: n = 6, ring = 28	16	62	60: ring = 17	1100	13,000
57: n = 2, ring = 24	150	2800			
 EtOCO-O -- F -- V -- O - L			 EtOCO- F - E - Nva - K		
58: n = 2, ring = 22	71	500	61: ring = 15	50,000	50,000
 EtOCO- C - F -V - O - L			 EtOCO- F - K - E - L		
59: ring = 18	690	5,200	62: ring = 12	1,200	5,300

In summary, there appears to be good reason to believe that the tetrapeptide (**1**) does bind in the manner of full length peptides and in the region corresponding to $HA_{309-312}$. There seems to be a certain degree of flexibility in the class II binding site. These SAR and their implications with regard to interpreting the DR/HA co-crystal for structure based design were used in our efforts to develop peptidomimetic ligands. This work will be described in coming reports.[17]

References and Notes

1. e-Mail address brian_jones@merck.com
2. Germain, R. N. *Cell*, **1994**, *76*, 287.
3. See Sloan, V. S.; Cameron, P.; Porter, G.; Gammon, M.; Amaya, M.; Mellins, E.; Zaller, D. M. *Nature (London)* **1995**, *375*, 802 and references cited.
4. Stern, L. J.; Brown, J. H.; Jardetzky, T. S.; Gorga, J. C.; Urban, R. G.; Strominger, J. L.; Wiley, D. C *Nature (London)* **1994**, *368*, 215.
5. Wucherpfennig, K. W.; Strominger, J. L. *J. Exp. Med.* **1995**, *181*, 1597. Of particular interest is the epidemiologic association of DR4 & DR1 expression with susceptibility to rheumatoid arthritis.
6. R. Cummings, presented at McMaster University, May 1996. Manuscript in preparation.
7. A fluorescence readout assay was employed that measures the ability of test peptides to inhibit binding of a biotinylated rat myelin basic protein peptide ($RMBP_{90-102}$) to purified DRB1*0101 (or DRB1*0401). A description of the protocol is provided in ref 10. We give IC_{50}s measured at two time points in this dynamic assay. This issue will not be discussed here since it does not influence the discussion. The data is shown in part to provide a second comparative data point in each case but also for reference with regard to upcoming reports wherein the data is of more significance and will be discussed in detail.

8. *Selective Immunosuppression: Basic Concepts and Clinical Applications* Adorini, L. Ed.; Karger; Basel.
9. Many studies have revealed this. For example see Hammer, J.; Belunis, C.; Bolin, D.; Papadopoulos, J.; Walsky, R.; Higelin, J.; Danho, W.; Sinigalia, F.; Nagy, Z. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4456.
10. Hill, C. M.; Liu, A.; Marshall, K. W.; Mayer, J.; Jorgensen, B.; Yuan, B.; Cubbon, R. M.; Nichols, E. A.; Wicker, L. S.; Rothbard, J. B. *J. Immunology* **1994**, *152*, 2890
11. In comparing structural motifs for promiscuous peptides binding to a number of DR alleles O'Sullivan[14] examined the effect of individual substitutions in the context of $HA_{307-319}$ and $TT_{830-843}$. From this data one can readily extract an 'optimum' tetrapeptide sequence of FV(K,R,S,A)(L,A) [$HA_{309-312}$ data] or FKIA [$TT_{831-834}$ data] corresponding to the proposed frame. That this is very similar to our tetrapeptides is quite compelling. Indeed, some of the peptides we examined are encompassed by this genus and have relevant IC_{50}s EtOCO-FVAL-NH_2 (28 nM), EtOCO-FVKL-NH_2 (32 nM) & free amino acid FVKL (850 nM). In addition, the frequent appearance of Arg at the P2 site of DRB1*0401 binding peptides has been noted (reference 9). One could ask whether a similar short peptide could be found in the HA T_{314}-L_{317} frame, i.e., anchored at L_{317} (the only other residue to significantly penetrate the binding groove, though considered a 'minor' anchor) and hitting the 314 minor pocket. A similar analysis of O'Sullivan's data suggested the sequences SQRL or ASSV. Neither EtOCO-SQRL-NH_2 nor EtOCO-ASSV-NH_2 showed any proclivity to bind.
12. In this communication we focus on the P1 and P4 positions since, according to this hypothesis, we would expect them to be the major side-chain determinants. Information on P2 and P3 variation will be found in ref 6.
13. **50** also correlates with full length peptide SAR - see ref 14 residue 834 of $TT_{830-843}$.
14. O'Sullivan, D.; Arrhenius, T.; Sidney, J.; Del Guercio, M-F.; Albertson, M.; Wall, M.; Oseroff, C.; Southwood, S.; Colon, S. M.; Gaeta F. C. A.; Sette, A. *J. Immunol.* **1991**, *147*, 2663.
15. This is also consistent with known SAR for DRB*0401 binding peptides (see Sette, A. et al. *J. Immunology*, **1993**, *151*, 3163 and McNicholl, J. M. et al. *J. Immunology*, **1995**, *155*, 1951). Both Pheβ13 and Argβ71 of DR1 may influence this but the DR4 polymorphism (Hisβ13 and Lysβ71 respectively) suggests that the selectivity is more likely based on the β13 interaction. The crystallographic distance between the HA_{312} γ-carbon and both the Pheβ13 2-position and the nearest Argβ71 N-atom is 3.6Å.
16. We confirmed the presence of the appropriate cycle in an HMBC NMR experiment.
17. During preparation of this manuscript a report appeared describing a hexapeptide ligand for DR4 and preliminary modification thereof (Hanson, G. J. et al. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1931). Some of our findings regarding peptidomimetic modification of these tetrapeptides were presented at the 25th National Medicinal Chemistry Symposium, Ann Arbor, MI, June 1996 (Jones, A. B. et al and Adams, A. D. et al.) and at the 14th International Medicinal Chemistry Symposium, Maastricht, Netherlands, September 1996 (Jones, A. B. et al.).

(*Received in USA* 18 *October* 1996; *accepted* 22 *November* 1996)