Bioorganic & Medicinal Chemistry Letters, Vol.3, No.12, pp. 2851-2856, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00

STRUCTURE-DIRECTED DISCOVERY OF AN INHIBITOR OF THE BINDING OF HIV GP120 TO THE CD4 RECEPTOR

Richard L. Jarvest,[a*] Ann L. Breen,[b] Colin M. Edge,[a] Margery A. Chaikin,[b]
L. John Jennings,[a] Alemseged Truneh,[b] Raymond W. Sweet[b] and Robert P. Hertzberg[b]

SmithKline Beecham Pharmaceuticals
a. Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ, U.K.
b. King of Prussia, Pennsylvania 19406, USA

(*Received in USA* 13 *August* 1993; *accepted* 12 *October* 1993)

Abstract Structure-directed screening identified the diphenyl lactone (**1**) as a specific inhibitor of the interaction of HIV gp120 with CD4. The compound bound selectively to gp120 in a time-dependent and irreversible fashion.

The ability of the human immunodeficiency virus (HIV) to infect the majority of its susceptible cell types is dependent on the high affinity interaction of the viral envelope glycoprotein gp120 with the cellular receptor CD4. The interaction between these proteins has been extensively studied[1,2] because of its potential as a target for specific therapeutic intervention in viral replication. Our aim was to find a low molecular weight inhibitor of the gp120-CD4 interaction which selectively binds to the viral protein gp120. Here we describe the discovery of such a compound by a structure-directed screening strategy.

The gp120 binding site on CD4 is localised on its N-terminal domain (D1), whose structure was recently determined in the form of D1D2, a soluble two-domain fragment of CD4.[3,4] D1D2, as well as soluble D1 fragments,[5,6] retain full binding affinity for gp120. Extensive mutagenesis of the terminal domain of CD4 has been carried out to define the residues involved in the interaction. In initial studies, the major determinants of the gp120 binding site were found to lie within the amino acids 40 - 60 (summarised in reference 4). In particular, two solvent-exposed residues, phenylalanine 43 and arginine 59, appear to be primary interaction sites. Phenylalanine 43 is the dominant determinant and, for retention of high gp120 binding affinity, Arg-59 can only be replaced by lysine, indicating that an interaction with a basic (positively charged) function may be necessary at this site (K. Deen *et al*, unpublished observations and ref 7). These two residues are closely located on the surface of the X-ray structure of D1D2, suggesting the possibility of mimicing the gp120 binding site with a small molecule containing an aromatic ring and a basic functionality.

The location of the Phe-43 and Arg-59 residues in one of the X-ray structures[4] of the D1D2 fragment of CD4 are shown in Figure 1.

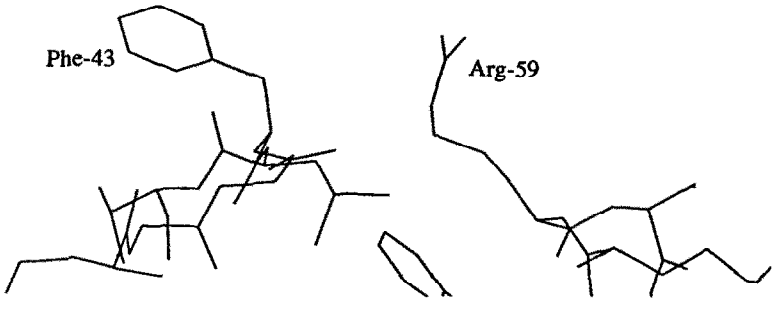


Figure 1. View of Phe-43 and Arg-59 on the surface of the X-ray structure of D1D2.[4]

The bond from Cβ of phenylalanine 43 to the phenyl ring is disposed roughly in-line with the guanidine moiety of arginine 59 and the distance from the Phe-43 Cβ to the Nε of Arg-59 is 4.7Å. The distances involved are such that we envisaged mimicing the relationship of these residues with a cyclic system which could function as a spacer unit to establish the correct distance and orientation of an aromatic group and a basic functionality as ring substituents.

A variety of cyclic systems substituted with a phenyl or benzyl substituent and with a basic side chain were modelled and fitted to the phenyl ring of Phe-43 and the guanidine of Arg-59. Some compounds were synthesised on this basis but the examples prepared showed no inhibition of gp120-CD4 binding.[8,9] However, the side chains of both Phe-43 and Arg-59 are highly solvent exposed and likely to adopt many orientations; thus the conformation in the complex with gp120 may be significantly different from that observed in this unbound structure. To allow for greater conformational change of these side chains, the parameters around the aryl-cyclic spacer-basic side chain were loosened and the model was used to search the SmithKline Beecham compound file in MACCS.[10] The search basis was a 5 or 6 membered ring system (carbocyclic or heterocyclic and having any degree of unsaturation from fully saturated to aromatic) which was 1,3 or 1,4 substituted with a phenyl or benzyl moiety and with a side chain containing a basic nitrogen atom at a distance of between 1 and 6 atoms from the ring.

About 80 available compounds selected in this way were assayed in an ELISA measuring inhibition of binding of soluble CD4 (sCD4) to immobilised recombinant gp120.[11,12] Of these compounds, it it was found that the lactone **1** (as its hydrochloride salt)[13,14] had an IC_{50} of 35 μM for inhibition of the gp120-sCD4 binding interaction (Figure 2). In addition, compound **1** was also active in an HIV viral syncytium assay[15,16] with an IC_{50} of 100 μM although there was some visible evidence of toxicity in this assay at 250 μM.

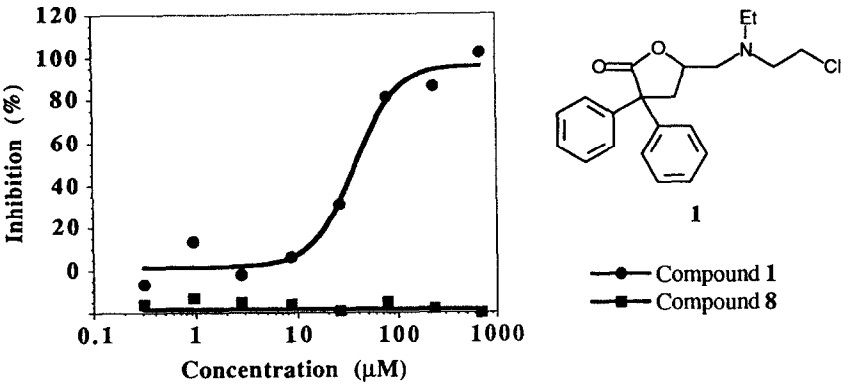
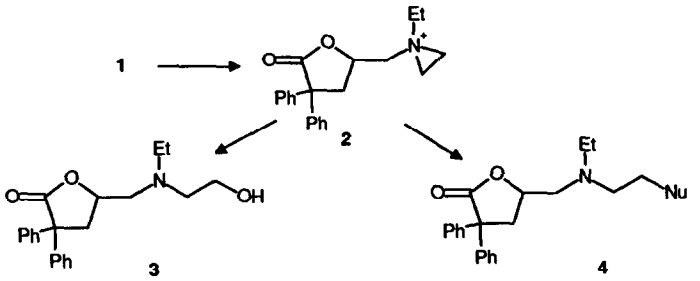
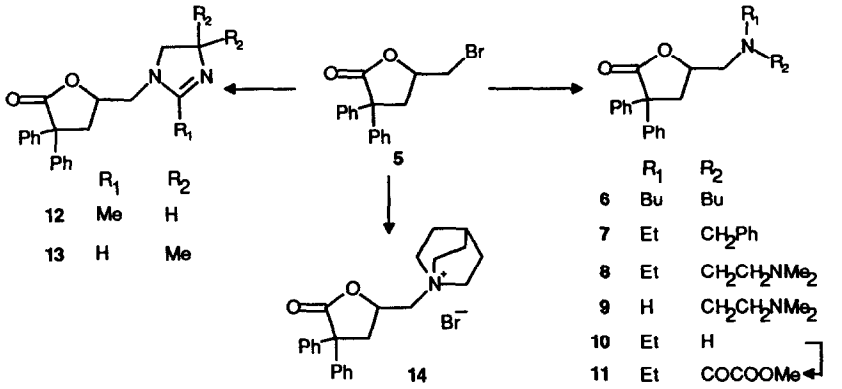


Figure 2. Inhibition of sCD4 binding to gp120 in the ELISA assay.

Modelling of **1** indicated the the *trans* phenyl ring could have a suitable geometry relative to the amine to mimic the Phe-43/Arg-59 geometry. However, due to its reactivity, the 2-chloroethylamine side chain could be acting in a number of ways. In principle there could be a reversible interaction with the 2-chloroethylamine salt itself, with the quaternary aziridinium intermediate **2** (rapidly formed from the deprotonated amine salt) or with the 2-hydroxyethylamine hydrolysis product **3**. Alternatively there could be covalent binding to protein as in **4**. A number of analogues of **1** were synthesised to explore these possibilities, whilst the hydroxy compound **3** was already available.[17]



The amines **6** to **10**, **12** and **13** were prepared by reaction of the bromo lactone **5**[18] with the neat amine at 110 to 130°C. The compounds include examples of alternative functional groups on the 2-ethyl position (**8** and **9**), examples where the chloro group is replaced by other lipophilic groups (**6** and **7**) and examples where a more basic function is included to better resemble the strongly basic guanidine moiety (**12** and **13**). Reaction of **5** with ethylamine in ethanol gave **10** which was acylated with methyl oxalyl chloride to afford compound **11**. Reaction of **5** with quinuclidine in diglyme afforded the quaternary salt **14**.



None of these compounds was active in the ELISA assay (Figure 2 and data not shown) although compound **8** was able to attenuate the inhibition shown by **1** (see below). This suggested that the 2-chloroethyl function played a critical role in the activity of **1**, possibly by alkylation of either CD4 or gp120 and its interaction with these proteins was investigated further.

The ability of **1** to block the CD4-gp120 interaction was confirmed in a RIA measuring binding of ^{125}I-sCD4 to gp120.[19] The IC_{50} of **1** in this assay was 30 μM. In a similar RIA format, the effect of compound **1** on sCD4 was assessed with a series of anti-CD4 monoclonal antibodies. These antibodies recognise different sites on CD4: L120 binds in domain 4 and is noncompetitive with gp120;[20,21] L71 binds in domain 1 and is weakly competitive with gp120;[20] and Leu3a also binds in domain 1, is competitive with gp120 and, like gp120, requires Phe-43 for high affinity binding (K. Deen *et al*, unpublished observations). Compound **1** did not inhibit the binding of any of these antibodies, indicating that it does not disrupt the sCD4-gp120 interaction through an effect on sCD4.

The interaction with gp120 was examined in the ELISA by preincubation with compound **1** prior to addition of sCD4. Increasing concentrations of **1** were added to the the antibody-immobilised gp120 and incubated for varying time periods. The compound was then either removed or maintained, sCD4 was added and the assay completed in the usual way. The observed inhibition was time-dependent, increasing to a plateau at 1 hour pre-incubation (Figure 3). Moreover the addition of the wash-out step made little difference to the observed activity. Similarly, in the RIA, less than 10% of sCD4 binding ability was recovered 20 hours after

diluting out the compound. This indicates that **1** binds to gp120 and dissociates extremely slowly. In view of the moderately potent IC_{50} value and the 2-chloroethylamine structure of **1**, this suggests that **1** undergoes a covalent interaction with gp120, presumably alkylation.

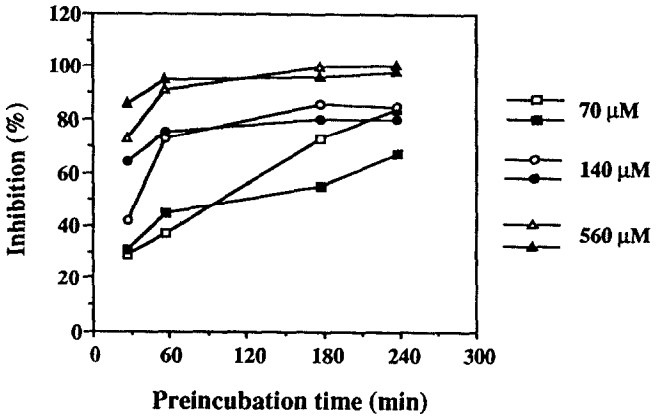


Figure 3. Time-dependent inhibition of sCD4 binding to gp120 by compound **1**. Compound was removed (open symbols) or maintained (closed symbols) prior to addition of sCD4.

The interaction of compound **1** with gp120 appeared to be specific and site-selective since the RIA was carried out in a high concentration of carrier milk protein and since the compound did not interfere with the binding of gp120 to the monoclonal antibody 178.1[22] which is non-competitive with CD4 binding. Preliminary results with other anti-gp120 monoclonal antibodies are also consistent with a site-selective interaction (J. Moore, personal communication). Competition assays with compound **8** suggest that this specificity is conferred by discrete structural features of **1**. Thus, pre-incubation of gp120 with **1** in the presence of varying concentrations of **8** for 90 minutes followed by addition of ^{125}I-sCD4 resulted in dose-dependent antagonism of inhibition by **1** (Figure 4).[23] Similar results were obtained in a parallel assay in which the preincubation time was 20 minutes. This suggests that the diphenyl-lactone-amino structure may be conferring the site-specificity of the interaction.

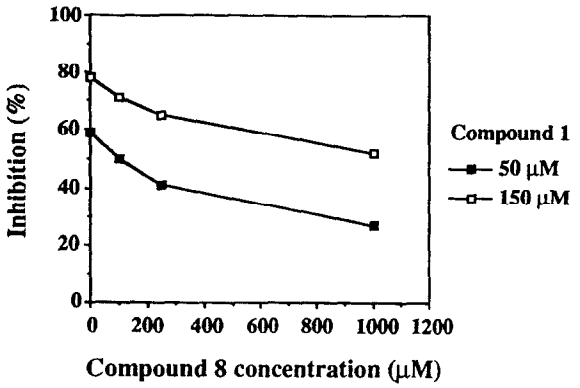
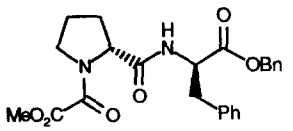


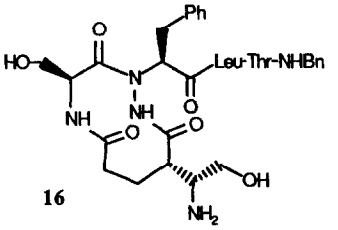
Figure 4. Antagonism of compound **1** inhibition of gp120-sCD4 binding by compound **8**.

Two other compounds based on the essential nature of Phe-43 have been described as inhibitors of the gp120-CD4 interaction. The phenylalanine-containing dipeptide, DD-CPF (**15**) was reported to inhibit this

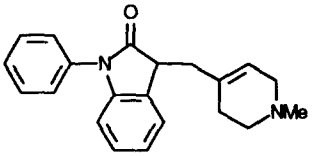
interaction at high concentrations, apparently by irreversible binding to gp120.[24] In our ELISA, the IC_{50} for **15** was 300μM, an order of magnitude higher than for **1**. Subsequently, it was found that **15** non-selectively inhibits a number of enveloped viruses through disruption of the integrity of the virus particles.[25] The second compound, the cyclic hexapeptide analogue **16**, was designed to mimic the loop structure on which Phe-43 is located.[26] This rather complex molecule was reported to inhibit the CD4-gp120 interaction with an estimated K_d of between 4 and 20 μM. However, in our ELISA, this compound was minimally active with an IC_{50} > 100 μM.



15



16



17

Interestingly, a series of anti-HIV compounds was recently described that resulted from the identification of **17** as a screening lead.[27] The authors suggest that the compounds may interact either with gp120 or CD4. Clearly **1** and **17** have some similarity: a 1,1 diphenyl moiety α to a carbonyl of a 5-membered ring lactam or lactone (although in **17** one phenyl is fused to the ring) and an amino containing side chain γ to the phenyl substituent. The reported SAR of **17** indicates that the phenyl ring is essential for high activity. Whether **17** also binds to gp120 has yet to be determined.

Acknowledgements We thank J.J. Leary and C. Anders for determination of anti-HIV activity and A.D. Patil for structure confirmation of compound **1**.

References and notes

1. Capon, D.J.; Ward, R.H.R. *Annu. Rev. Immunol.*, **1991**, *9*, 649.
2. Sweet, R.W.; Truneh, A.; Hendrickson, W.A. *Current Opinion in Biotechnology*, **1991**, *2*, 622.
3. Wang, J.; Yan,Y.; Garrett, T.P.J.; Liu, J.; Rodgers, D.W.; Garlick, R.L.; Tarr, G.E.; Husain, Y.; Reinherz, E.L.; Harrison, S.C. *Nature*, **1990**, *348*, 411.
4. Ryu, S.-E.; Kwong, P.D.; Truneh, A.; Porter, T.G.; Arthos, J.; Rosenberg, M.; Dai, X.; Xuong, N.; Axel, R.; Sweet, R.W.; Hendrickson, W.A. *Nature*, **1990**, *348*, 418.
5. Chao, B.H.; Costopoulos, D.S.; Curiel, T.; Bertonis, J.M.; Chisholm, P.; Williams, C.; Schooley, R.T.; Rosa, J.J.; Fisher, R.A.; Maraganore, J.M. *J. Biol. Chem.*, **1989**, *264*, 5182.
6. Arthos J.; Deen, K.C.; Chaikin, M.A.; Fornwald, J.A.; Sattentau, Q.J.; Clapham, P.R.; Weiss, R.A.; McDougal, J.S.; Pietropaolo, C.; Axel, R.; Truneh, A.; Maddon, P.J.; Sweet, R.W. *Cell*, **1989**, *57*, 469.
7. Moebius, U.; Clayton, L.; Abraham, S.; Harrison, S.C.; Reinherz, E.L. *J. Exp. Med.*, **1992**, *176*, 507.
8. Jarvest, R.L.; Readshaw, S.A. *Synthesis*, **1992** 962.
9. Jarvest, R.L.; Marshall, J.E. *J. Heterocyclic Chem*, **1992**, *29*, 1401.
10. Version 1.5; Molecular Design Limited, San Leandro, CA, USA.
11. Hertzberg, R.P.; Johnson, R.K.; Caranfa, M.J.; Meyers, C.C.; MacKenzie, L.; Breen, A.L.; Tomaszek, T.A.; Meek, T.D.; Debouck, C.; Mai, S.H.; Culp, J.S.; Francis, T.A. *New Drugs from Natural Sources*; Coombes J.D., Ed.; IBC Technical Services Ltd, London, 1992, pp 46-58.
12. An Immulon-2 microtitre plate was incubated overnight at 4°C with 2 μg goat anti-mouse IgG in 100 μl PBS (0.01 M phosphate buffer, 0.15 M NaCl), pH 6.8, 0.005% glutaraldehyde. The solution was then removed and the plate washed with pH 7.4 PBS (x 3). Mouse anti-gp120 antibody 178.1[22] (0.2 μg) was added in 100 μl PBS-Tween-BSA (0.5% BSA, 0.05% Tween 20) and incubated overnight at room temperature (RT). The antibody was removed from the plate, the plate was washed, and gp 120[28] (0.1

μg) in 100 μl PBS-Tween-BSA was added and incubated at 37°C for 1 h. The plate was washed and blocked with PBS-Tween-BSA for 1 h at RT, washed again, and then 50 μl of sCD4 (100 ng/ml)[29] plus 50 μl test substance was added and incubated for 1 h ar RT.[29] The plate was washed and 100 μl of peroxidase-conjugated rabbit anti-sCD4 IgG added, and incubated for 1 h at RT. The bound sCD4 was then quantitated using a colourimetric enzyme reaction.[30]

13. Arnold R.T.; Ullyot G.E. US Patent 2,650,231 (to Smith, Kline and French Laboratories) **1953.**
14. The identity of this sample was confirmed by ^{1}H nmr and mass spectroscopy.
15. Truneh, A.; Frescatore, R.L.; Thiem, P.; Leary, J.J.; Rosenberg, M.; Hanna, N.; Sweet, R.; Bugelski, P.J. *J. Cell. Immunol.*, **1990**, *131*, 98.
16. As described[15] except syncytia were counted at 24 h. In this assay the IC_{50} for sCD4 was 5 nM.
17. Craig, P.N.; Witt, I.H.; Macko, E.; Dacanay, J.G; Fellows, E.J.; Ullyot, G.E. *J. Am. Chem. Soc.*, **1951**, *73*, 1339.
18. Craig, P.N.; Witt, I.H. *J. Am. Chem. Soc.*, **1950**, *72*, 4925.
19. For the CD4 antibody RIA, serial dilutions of **1** in 10% DMSO were mixed with ^{125}I-sCD4 (1 nM, about 200 μC/nmol) in RIA buffer (PBS containing 0.25% dry milk plus 0.05% NP40) and added to microtitre plates which had been precoated sequentially with goat anti-mouse IgG and the CD4 monoclonal antibody and then blocked with 0.1% gelatin. The plates were incubated for 3 h at RT, washed with PBS and the wells counted. The IC_{50} for sCD4 was 1-5 nM.
For the gp120 RIA, serial dilutions of **1** were mixed with ^{125}I-sCD4 as above and then mixed with gp120 to give a final concentration of 1 nM of gp120 and sCD4. The samples were processed as above except that the gp120 antibody 178.1 was used in the capture step. The IC_{50} for sCD4 was 5 nM.
20. Truneh, A.; Buck, D.; Cassatt, D.R.; Juszczak, R.; Kassis, S.; Ryu, S.-E.; Healey, D.; Sweet, R.; Sattentau, Q. *J. Biol. Chem.*, **1991**, *266*, 5992.
21. Healey, D.; Dianda, L.; Moore, J.P.; McDougal, J.S.; Moore M.J.; Estess, P.; Buck, D.; Kwong, P.D.; Beverly, P.C.L.; Sattentau, Q.J. *J. Exp. Med.*, **1990**, *172*, 1233.
22. Thiriart, C.; Francotte, M.; Cohen, J.; Collignon, C.; Delers, A.; Kummert, S.; Molitor, C.; Gilles, D.; Roelants, P.; van Wijnedaele, F.; De Wilde, M.; Bruck, C. *J. Immunol.*, **1989**, *143*, 1832.
23. Compounds **1** (50 or 150 μM) and **8** (0-1 mM) were incubated at RT with gp120 for 20 or 90 min in RIA buffer. ^{125}I-sCD4 (1 nM) was then added and the incubation was continued for 1 h. The gp120/sCD4 complex was captured on protein A sepharose beads coated with the gp120 antibody 178.1, washed, and counted.
24. Finberg, R.W.; Diamond, D.C.; Mitchell, D.B.; Rosenstein, Y.; Soman, G.; Norman, T.C.; Schreiber, S.L.; Burakoff, S.J. *Science*, **1990**, *247*, 287.
25. Moore, J.P.; Simpson, G.; McKeating, J.A.; Burakoff, S.J.; Schreiber, S.L.; Weiss, R.A. *Virology*, **1992**, *188*, 537.
26. Chen, S.; Chrusciel, R.A.; Nakanishi, H.; Raktabutr, A.; Johnson, M.E.; Sato, A.; Weiner, D.; Hoxie, J.; Saragovi, H.U. Greene, M.I.; Kahn, M. *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 5872.
27. Smallheer, J.M.; Otto, M.J.; Amaral-Ly, C.A.; Earl, R.A.; Myers, M.J.; Pennev, P.; Montefiori, D.C.; Wuonola, M.A. *Antiviral Chem. Chemother.*, **1993**, *4*, 27.
28. Culp, J.S.; Johansen, H.; Hellmig, B.; Beck, J.; Matthews, T.J.; Delers, A.; Rosenberg, M. *Bio/Technology*, **1991**, *9*, 173.
29. Deen, K.C.; McDougal, J.S.; Inacker, R.; Folena-Wasserman, G.; Arthos, J.; Rosenberg, J.; Maddon, P.J.; Axel, R.; Sweet, R.W. *Nature*, **1988**, *331*, 82.
30. Tijssen, T *Practice and Theory of Enzyme Immunoassays*; Burdon, R.H. Knippenberg, P.H., Ed.; Elsevier Science Publishers B.V., the Netherlands, 1985, pp361-365.