Identification of Residues in CD6 Which Are Critical for Ligand Binding

Dale L. Bodian, John E. Skonier, Michael A. Bowen, Michael Neubauer, Anthony W. Siadak, Alejandro Aruffo,[‡] and Jürgen Bajorath*,[‡]

Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, Washington 98121, and Department of Biological Structure, University of Washington, Seattle, Washington 98195

Received October 11, 1996; Revised Manuscript Received December 12, 1996[⊗]

ABSTRACT: CD6 is a member of the scavenger receptor cysteine rich protein superfamily (SRCRSF). This family includes many cell surface proteins whose three-dimensional structures and functions are presently not well understood. The extracellular region of CD6 includes 3 SRCR domains. The membrane proximal SRCR domain specifically binds the activated leukocyte cell adhesion molecule (ALCAM), a CD6 ligand belonging to the immunoglobulin superfamily. CD6−ALCAM interactions mediate immune cell adhesion and are implicated in T cell maturation and the regulation of T cell function. On the basis of SRCRSF sequence comparison, a mutagenesis analysis of the membrane proximal SRCR domain of CD6 (CD6D3) has been carried out. Fifteen mutants were characterized. Three CD6 residues were identified in a region of low sequence conservation which, when mutated, abolish ligand binding but not the binding to a panel of conformationally sensitive anti-CD6 mAbs. This study provides the first analysis of residues critical for ligand binding to a member of the SRCRSF.

SRCR domains[1] are found in a variety of cell surface molecules and some secreted proteins. Members of this protein superfamily include, for example, the macrophage scavenger receptor type I, the speract receptor, the complement factor I, and the leukocyte antigens CD5, M130, and WC1 (Resnick et al., 1994). In contrast to the IgSF (Williams & Barclay, 1988), the three-dimensional structures, functions, and ligands of SRCRSF proteins are presently poorly characterized. No information is currently available about the structural or functional roles of residues comprising SRCR domains.

CD6, a 100−130 kDa cell surface protein expressed on mature T cells, a subpopulation of B cells, and some cells in the brain (Reinherz et al., 1982; Morimoto et al., 1988; Aruffo et al., 1991) is a member of the SRCRSF and has been implicated as a costimulatory molecule on T cells (Vollger et al., 1987; Morimoto et al., 1988; Gangemi et al., 1989). The extracellular region of CD6 contains 3 SRCR domains and a ∼30-residue stalk (Aruffo et al., 1991). This is followed by a transmembrane domain and an alternatively spliced cytoplasmic region (Robinson et al., 1995) which becomes phosphorylated in activated T cells (Wee et al., 1993).

Recently, a human CD6 ligand, ALCAM, was cloned (Bowen et al., 1995). In the immune system, ALCAM is expressed on activated T and B cells, monocytes, and TE cells (Wee et al., 1994; Patel et al., 1995; Bowen et al., 1996). The CD6-ALCAM interaction mediates thymocyte−TE cell adhesion (Bowen et al., 1995) and has thus been implicated in T cell development. ALCAM is a member of the IgSF, and its extracellular region includes five Ig-like domains (Bowen et al., 1995). The interaction between CD6 and its ligand has been studied in some detail. Using soluble recombinant proteins, it was found that the membrane proximal SRCR domain of CD6 (Whitney et al., 1995) binds to the N-terminal Ig V-like domain of ALCAM with 1:1 stoichiometry (Bowen et al., 1996). Several ALCAM residues have been identified as important for the interaction with CD6 (Skonier et al., 1996). These residues map to the predicted A′GFCC′C″ face of the N-terminal Ig V-like domain (Skonier et al., 1996).

Here we report on the identification of residues in human CD6 which are critical for ALCAM binding. Sequences of different SRCR domains were compared. On the basis of this comparison a systematic mutagenesis study was undertaken. The binding of 15 mutants of the membrane proximal SRCR domain of CD6 (CD6D3) to a panel of anti-CD6D3 mAbs and ALCAM was tested. The amino (N)-terminal half of many SRCR domains displays striking residue conservation, whereas the carboxy (C)-terminal half is much more variable in sequence. Three residues in the C-terminal half of CD6D3 were identified as critical for ligand binding.

MATERIALS AND METHODS

Sequence Analysis. Proteins which include SRCR (-like) domains were identified by BLAST searches in a compilation of the SwissProt, PIR, and GenPept databases. SRCR domain sequences were initially aligned using PileUp (Feng & Doolittle, 1987) implemented in the GCG program

* Correspondence to BMS-PRI, 3005 First Ave., Seattle, WA 98121; Tel, (206) 727−3612; Fax, (206) 727−3602; E-mail, bajorath@protos.bms.com.

‡ University of Washington and BMS-PRI.

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1997.

[1] Abbreviations: ALCAM, activated leukocyte cell adhesion molecule; CD6D3, membrane proximal extracellular domain of human CD6; CD6D3S, membrane proximal extracellular domain of human CD6 including the stalk region; ELISA, enzyme linked immunosorbant assay; HRP, horseradish peroxidase; Ig, immunoglobulin; IgG, immunoglobulin G; IgSF, immunoglobulin superfamily; Ig-V domain, immunoglobulin variable domain; mAb, monoclonal antibody; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; RT, room temperature; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SRCR domain, scavenger receptor cysteine rich domain; SRCRSF, scavenger receptor cysteine rich protein superfamily; TE, thymic epithelium; WT, wild type.

S0006-2960(96)02560-3 CCC: $14.00

package (Genetics Computer Corporation, Madison, WI). For proteins with multiple SRCR domains, only a single domain was aligned except for CD5 and CD6 sequences, for which all domains were considered. The alignment was manually modified using CAMELEON (Oxford Molecular Ltd., Oxford, U.K.) to ensure consistent alignment of cysteine positions and the appropriate placement of gaps. Gaps were introduced in regions of high sequence variability. In modifying the alignment, residue character (e.g., hydrophobic, charged/polar etc.) was conserved as much as possible. Sequences included in the alignment were X60992, U12434, X78985, D10728, M97195, X63723, S76313, U32681, U37438, S78981, U20652, S78869, and Z22968 from GenBank, P19238 from SwissProt, and A57190 from PIR.

Monoclonal Antibodies. Murine mAbs were raised against CD6−Ig fusion protein and selected for their ability to bind to CD6D3. Four mAbs, A, B, C, and D, were used to assess the conformational integrity of CD6D3S mutant proteins. These mAbs do not show reactivity in Western blots and thus recognize conformational epitopes. These mAbs also block ALCAM binding to CD6 suggesting that, while the epitopes of these mAbs are unknown, they bind at least proximal to the ligand binding site.

Site Directed Mutagenesis. Point mutations and diagnostic restriction sites were introduced in CD6D3 including the stalk region (CD6D3S) and inserted into the plasmids CDM7B- (Hollenbaugh et al, 1995) or PD19 (unpublished). Both vectors yield fusion proteins including the constant and hinge domains of human IgG_1. Mutants were introduced into CDM7B- with the QuikChange mutagenesis kit (Stratagene) following the manufacturer's directions. Mutated plasmids were transformed by electroporation into competent MC1061 cells, followed by AmpTet selection. Other mutants were generated by the QuikChange procedure on the PD19 vector containing the CD6D3S insert, followed by heat shock transformation by into supercompetent XL-1 blue cells (Stratagene) and Amp selection. Mutants were also generated by overlap extension PCR as described (Bajorath et al., 1995). The sequence of the entire CD6D3S coding region was confirmed for each mutant by dideoxy sequencing with the Thermosequenase cycle sequencing kit (Amersham) following the manufacturer's directions. mAb and ligand binding of CD6 wild type proteins produced by the two different vectors were indistinguishable by ELISA.

Protein Expression. Mutant fusion proteins were produced by transient expression in COS cells as described (Bajorath et al., 1995). The expression of fusion proteins was monitored by ELISA as described below for determination of protein concentration. In addition, purified mutant proteins were compared to wild type by SDS-PAGE and native PAGE to ensure the correct size of expressed proteins.

Protein Purification. Tissue culture supernatant (7 mL) was incubated with 100 μL of a 50% suspension of protein A-sepharose beads overnight at 4 °C with constant stirring. Samples were washed twice with PBS, then transferred to a Millipore ultra free MC (0.22 μm). Excess PBS was removed by microcentrifugation. Proteins were eluted in 50 μL of 0.1 M sodium citrate, pH 3.0, following a 5 min incubation at RT.

Determination of Protein Concentration by ELISA. Immulon-2 96-well plates (Dynatech Labs) were coated with goat anti-human IgG (Jackson ImmunoResearch), diluted to 1:1000 with carbonate buffer, by incubation at 4 °C overnight. Wells were washed four times with PBS/.05% Tween, then incubated in 1X EIA diluent (Genetic Systems) for 1 h at RT, and washed again. Filtered tissue culture supernatants containing the mutant fusion proteins were serially diluted with EIA diluent, then incubated on the plates for 1 h at RT. An Ig fusion protein containing the 3 N-terminal domains of ALCAM (Bowen et al., 1996) at known concentration was used as a standard. After washing, HRP-conjugated donkey anti-human IgG (Jackson ImmunoResearch), diluted 1:2000 in EIA, was added to each well, and plates were incubated for 1 h at RT. Plates were then developed by the addition of 70 μL TMB solution (Genetic Systems) diluted 1:100 in citrate-buffered substrate solution (Genetic Systems). Reactions were stopped after 10 min at RT by the addition of 70 μL of 1 N H_2SO_4. The OD 450:630 ratio was determined for each well and averaged over triplicate samples.

mAb Binding by ELISA. Immulon-2 plates were coated with goat anti-human antibody, washed, blocked, and incubated with fusion protein as described above. Serial 2x dilutions of fusion proteins were made starting at 1 μg/mL. Hybridoma cell supernatants containing anti-CD6D3 mAbs were diluted at 1:50−1:70 in EIA diluent, added to plates, and incubated for 1 h at RT. Plates were washed in PBS/Tween, followed by incubation of goat anti-mouse IgG-HRP (Southern Biotechnology Associates) at 1:5000 for 1 h at RT. Plates were developed as above, except that the reactions were terminated after 15 min.

ALCAM Binding by ELISA. With the following exceptions, the assay format was the same as described for mAb binding. Serial 3x dilutions of fusion proteins were prepared starting at 0.5 μg/mL. Instead of adding anti-CD6 antibodies, a fusion protein consisting of full length ALCAM and mouse IgG was used at 0.05 μg/mL (Bowen et al., 1996). This ALCAM binding assay represents a high avidity assay format (Skonier et al., 1996).

RESULTS AND DISCUSSION

SRCRSF Sequence Analysis. In the absence of a three-dimensional SRCR domain structure, multiple sequence comparison was used to aid in the selection of CD6D3 residues for mutagenesis. BLAST searches identified 21 different proteins containing at least one SRCR (-like) domain. Previously, a subset of these proteins was used to classify SRCR domains into two families, A and B, based on the patterns of conserved cysteines and some other residues (Resnick et al., 1994). CD6D3 includes eight cysteines and belongs to the B set. Of the sequences found by BLAST, only B set SRCR domains with clear sequence similarity to CD6D3 were included in the comparison. Figure 1 shows the comparison of 27 SRCR domain sequences from 10 proteins and their homologs in different species. The comparison reveals that the N-terminal half of these SRCR domains, which include approximately 100 residues, displays an alternating pattern of conserved and variable sequence. With the exception of four cysteine positions and a few additional consensus residues, much greater sequence diversity is observed in the C-terminal half.

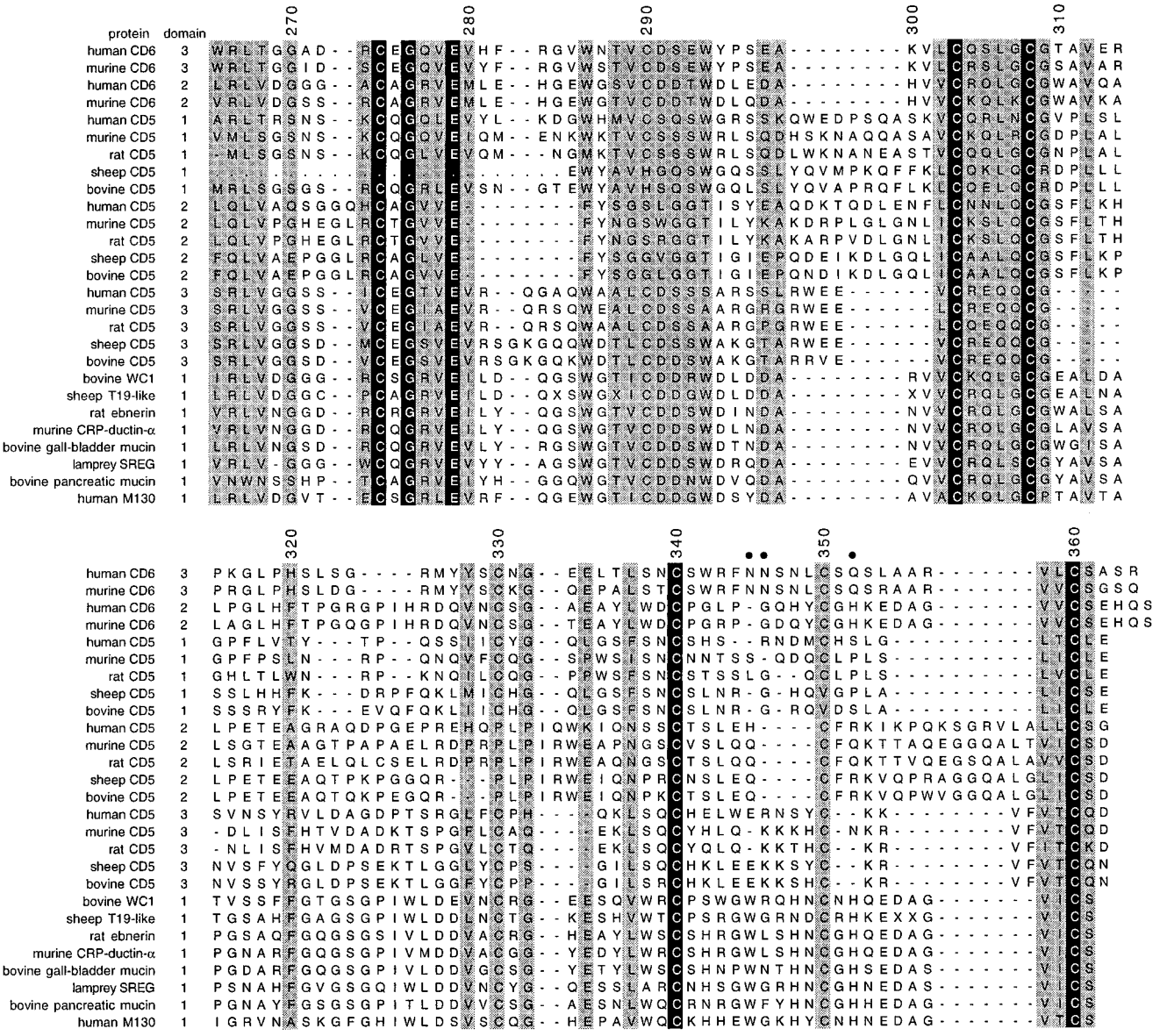

FIGURE 1: Comparison of SRCR domains. Sequences were selected and aligned as described in the text. Black shading indicates residues which are 100% conserved in all compared sequences, and grey shading highlights positions where at least 15 sequences include the same residues or conservative replacements. Sequence numbers are given for human CD6. The symbol "•" marks CD6D3 residues whose mutation disrupted ALCAM binding but not mAb binding. Gaps introduced for alignment are represented by dashes. Periods indicate unsequenced positions, and "X" denotes ambiguous sequence.

These findings suggested that residues which determine the specificity of an SRCR domain may reside in the C-terminal region.

Mutagenesis Approach. Residues were mutated throughout CD6D3, but the majority of targeted residues were in the C-terminal region. Since murine and human CD6 and ALCAM exhibit cross-species binding (Whitney et al., 1995), it is likely that at least some CD6D3 residues important for ligand binding are conserved. Several residues (R283, E293, S305, R314, Q352) were selected on this basis but residues not conserved in murine and human CD6 were also targeted. In addition, some residues highly conserved in all compared SRCR domains and thus expected to play a structural role were mutated. For example, E279 is the only non-cysteine residue that is rigorously conserved. Drastic mutations to R or other charged residues were introduced at all selected positions to significantly disrupt either the CD6D3 structure or CD6-ALCAM interactions.

Mutant Proteins and Binding Characteristics. Point mutants were generated as secreted CD6D3S−Ig fusion proteins. Fifteen of 17 mutants were expressed in sufficient quantities for further characterization (Table 1). E279R and R283E were expressed at levels which could not be detected by both ELISA and SDS-PAGE. Each of the 15 expressed mutant proteins was tested for binding to ALCAM and the panel of conformationally sensitive mAbs. Representative mAb binding experiments are shown in Figure 2. The ALCAM binding experiments are shown in Figure 3. The results of all experiments are summarized in Table 1. Five mutant proteins (Q277R, R314E, S353K, A355D, R357E) did not bind to ALCAM or any of the mAbs. E293R only bound to one mAb (Table 1). Three mutant proteins (A271R, V285E, N339D) bound to both mAbs and ALCAM like wild type CD6. The binding of other mutants, as discussed below, discriminated between mAbs and ALCAM.

Table 1: Binding of CD6D3 Mutants to ALCAM and mAbs[a]

mutant	mAb binding A	B	C	D	ALCAM binding
WT	+	+	+	+	+
A271R	+	+	+	+	+
Q277R	−	−	−	−	−
E279R	NA	NA	NA	NA	NA
R283E	NA	NA	NA	NA	NA
V285E	+	+	+	+	+
E293R	−	+	−	−	−
S321K	+	+	+	+	+
N339D	+	+	+	+	+
N339D/N345D	+	+	+	+	−
N345D	+	+	+	+	−
N345Q	+/−	+/−	+/−	+/−	−
N346K	+	+	+	+	−
Q352R	+	+	+	+	−
S353K	−	−	−	−	−
A355D	−	−	−	−	−
R357E	−	−	−	−	−
S363K	+/−	+/−	+/−	+/−	+/−

[a] NA, not expressed at detectable levels; +, binding comparable to wild type; +/−, reduced binding; −, no detectable binding.

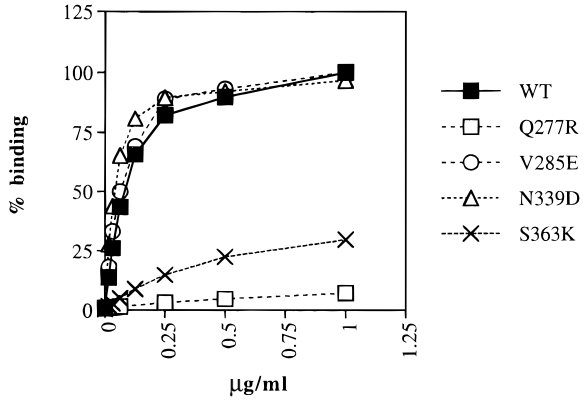

FIGURE 2: Binding of CD6D3 mutants to mAbs. Representative ELISA experiments are shown for the binding of mutant proteins to mAb A. Mutants are designated according to Table 1. OD 450: 630 values were normalized relative to maximum OD values for wild type protein and reported as "% binding".

Residues Critical for Ligand Binding. CD6 residues were only considered critical for ligand binding if mutant proteins showed reduced binding to ALCAM but consistently bound at wild type levels to the conformationally sensitive mAbs. For example, mutant protein S363K bound only weakly to all four mAbs and to ALCAM and was therefore not considered. The limitation of this approach is that residues which are equally critical for ligand and mAb binding cannot be distinguished from residues whose mutation disrupts the three-dimensional structure of the protein.

Three CD6 mutant proteins, Q352R, N345D, and N346K, showed no detectable binding to ALCAM but bound at wild type levels to all four mAbs (Figure 3, Table 1). Thus, these residues, which include a potential N-linked glycosylation site at N345, were considered critical for ligand binding to CD6. The three residues map to the variable C-terminal region in CD6D3 (Figure 1) and are conserved in human and murine CD6, in accord with the observed cross-species CD6−ALCAM interaction (Whitney et al., 1995).

The importance of glycosylation for ALCAM binding was further explored. Mutation of residue N339 to D, a second potential glycosylation site in CD6D3, did not affect mAb or ALCAM binding (Table 1). Like N345D, the double

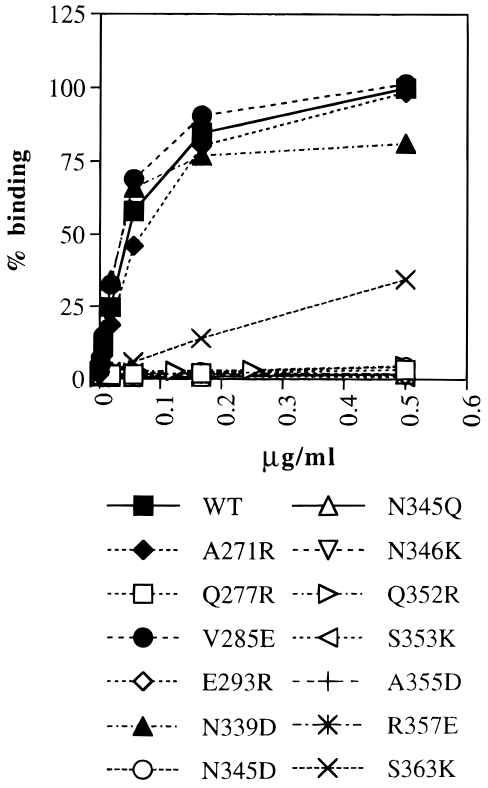

FIGURE 3: Binding of CD6D3 mutants to ALCAM. Filled symbols indicate mutants with wild type-like binding (corresponding to "+" in Table 1). "% binding" is reported according to Figure 2.

mutant N339D/N345D bound to all mAbs but not to ALCAM. When N345 was mutated to Q instead of D, mAb binding was consistently reduced, while no ALCAM binding was observed, indicating a structural perturbation. In aggregate, the results suggest that glycosylation at N345 but not N339 is critical for the integrity of the CD6−ALCAM interaction. Similar observations have been made, for example, in the case of CD2, but the molecular nature of these effects remains controversial (Davis & van der Merwe, 1996; Reinherz et al., 1996).

CONCLUSIONS

A site-directed mutagenesis approach, guided by multiple sequence comparison, was carried out to identify residues critical for ligand binding to CD6. The gross structural integrity of mutant proteins was assessed by binding to a panel of conformationally sensitive mAbs. Residues in the variable C-terminal half of CD6D3 were identified as critical for ALCAM binding. This is the first characterization of residues important for ligand binding to a member of the SRCRSF. Future studies will determine if similar sets of residues in other SRCR domains determine their binding specificity.

ACKNOWLEDGMENT

We thank Bill Bear for assistance with minipreps and oligonucleotide synthesis and Barb Dunning for help with the ELISA assays.

REFERENCES

Aruffo, A., Melnick, M. B., Linsley, P. S., & Seed, B. (1991) *J. Exp. Med. 174*, 949−952.

Bajorath, J., Chalupny, N. J., Marken, J. S., Siadak, A. W., Skonier, J., Gordon, M., Hollenbaugh, D., Noelle, R. J., Ochs, H. D., & Aruffo, A. (1995) *Biochemistry 34*, 1834−1844.

Bowen, M. A., Patel, D. D., Li, X., Modrell, B., Malacko, A. R., Wang, W.-C., Marquardt, H., Neubauer, M., Pesando, J. M., Francke, U., Haynes, B. F., & Aruffo, A. (1995) *J. Exp. Med. 181*, 2213−2220.

Bowen, M. A., Bajorath, J., Siadak, A. W., Modrell, B., Malacko, A. R., Marquardt, H., Nadler, S. G., & Aruffo, A. (1996) *J. Biol. Chem. 271*, 17390−17396.

Davis, S. J., & van der Merwe, P. A. (1996) *Science 273*, 1241−1242.

Feng, D.-F., & Doolittle, R. F. (1987) *J. Mol. Evol. 25*, 351−360.

Gangemi, R. M., Swack, J. A., Gaviria, D. M., & Romain, P. L. (1989) *J. Immunol. 143*, 2439−2447.

Hollenbaugh, D., Doutwright, J., McDonald, V., & Aruffo, A. (1995) *J. Immunol. Methods 188*, 1−7.

Morimoto, C., Rudd, C. E., Letvin, N. L., Hagan, M., & Schlossman, S. F. (1988) *J. Immunol. 140*, 2165−2170.

Patel, D. D., Wee, S.-F., Whichard, L. P., Bowen, M. A., Pesando, J. M., Aruffo, A., & Haynes, B. F. (1994) *J. Exp. Med. 181*, 1563−1568.

Reinherz, E. L., Meuer, S., Fitzgerald, K. A., Hussey, R. E., Levine, H., & Schlossman, S. F. (1982) *Cell 30*, 735−743.

Reinherz, E. L., Li, J., Smoylar, A., Wyss, D. F., Knoppers, M. H., Willis, K. J., Arulanandam, A. R. N., Choi, J. S., & Wagner, G. (1996) *Science 273*, 1242.

Resnick, D., Pearson, A., & Krieger, M. (1994) *Trends Biochem. Sci. 19*, 5−8.

Robinson, W. H., de Vegvar, H. E. N., Prohaska, S. S., Rhee, J. W., & Parnes, J. R. (1995) *Eur. J. Immunol. 25*, 2765−2769.

Skonier, J. E., Bowen, M. A., Emswiler, J., Aruffo, A., & Bajorath, J. (1996) *Biochemistry 35*, 14743−14748.

Vollger, L. W., Tuck, D. T., Springer, T. A., Haynes, B. F., & Singer, K. H. (1987) *J. Immunol. 138*, 358−363.

Wee, S.-F., Wang, W.-C., Farr, A. G., Nelson, A. J., Patel, D. D., Haynes, B. F., Linsley, P. S., & Aruffo, A. (1994) *Cell. Immunol. 158*, 353−364.

Whitney, G. S., Starling, G. C., Bowen, M. A., Modrell, B., Siadak, A. W., & Aruffo, A. (1995) *J. Biol. Chem. 270*, 18187−18190.

Williams, A. F., & Barclay, A. N. (1988) *Annu. Rev. Immunol. 6*, 381−406.

BI962560+