STRUCTURAL
BIOLOGY

Volume 77 (2021)

Supporting information for article:

Structural and functional analysis of the C-terminal region of *Streptococcus gordonii* SspB

Norbert Schormann, Sangeetha Purushotham, Joshua Mieher, Manisha Patel, Hui Wu and Champion Deivanayagam

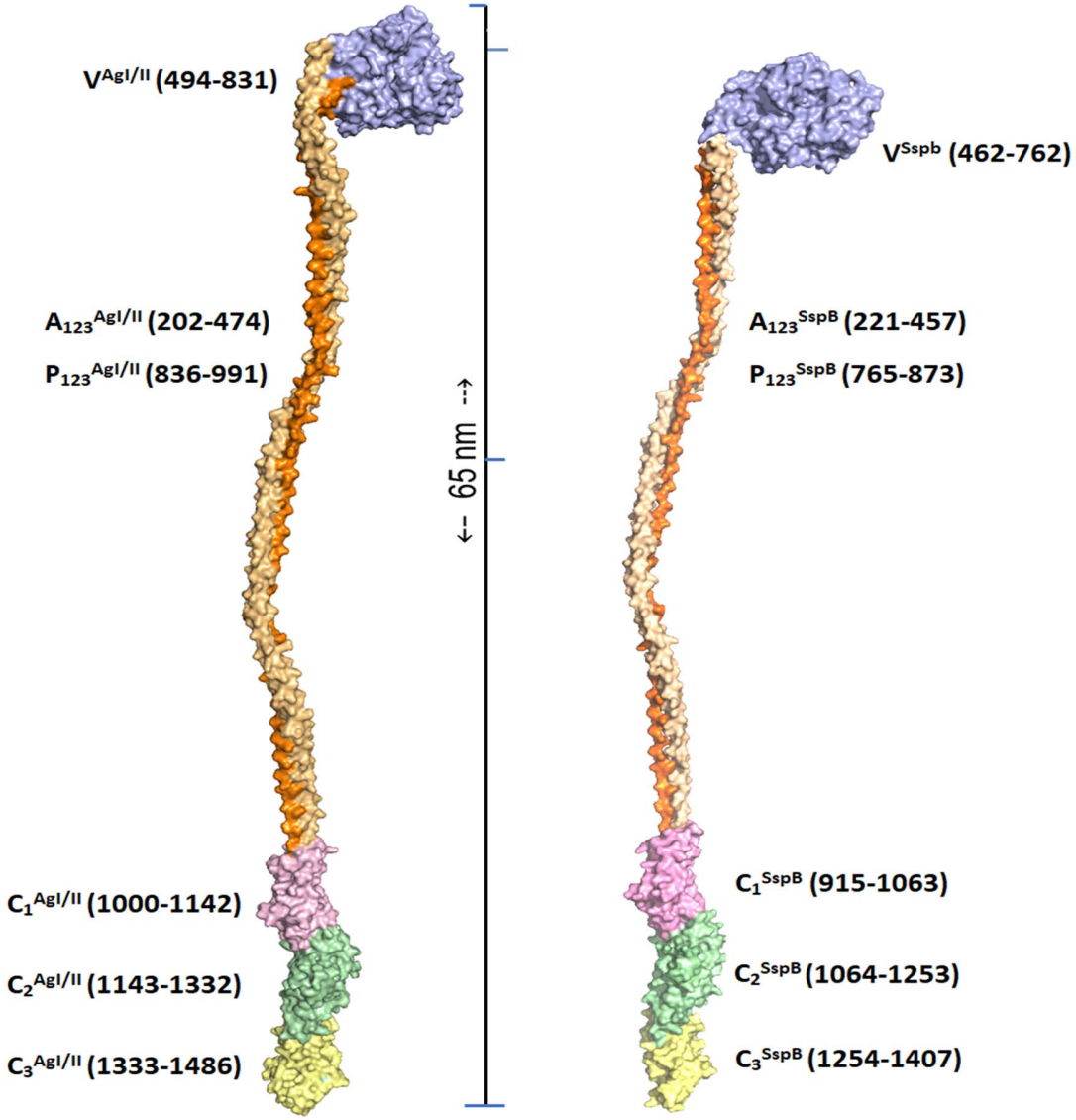

Figure S1 Models for FL^{SspB} and $FL^{AgI/II}$. The figure displays 3D models for $FL^{AgI/II}$ (residue range 202-1486) and FL^{SspB} (residue range 221-1407). Variable (V) domains, alanine-rich (A_{123}) repeats, proline-rich repeats (P_{123}) and C-terminal (C_{123}) domains are shown in surface representation in different colors, and they are labeled with their residue range indicated.

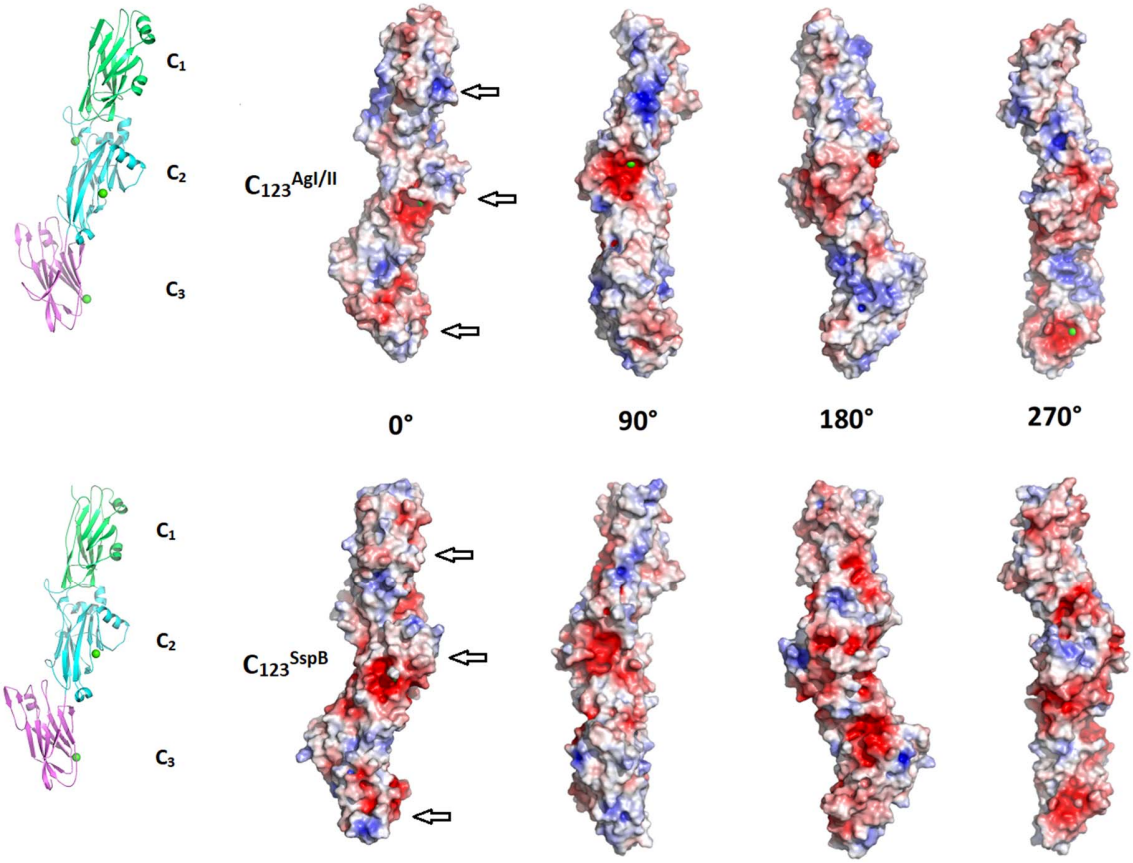

Figure S2 Molecular surface plots contoured by electrostatic potential for $C_{123}^{AgI/II}$ and C_{123}^{SspB}. The plots are shown in four orientations (three rotations about y-axis by 90º each). Areas of $C_{123}^{AgI/II}$ and C_{123}^{SspB} that are quite different in shape and charge in domains C_1, C_2 and C_3 are indicated by open arrows. For both superposed structures a ribbon diagram is provided with domains in different colors and labeled (C_1: green; C_2: cyan; C_3: violet). Contours for the electrostatic potential are: -5kT in red; +5kT in blue.

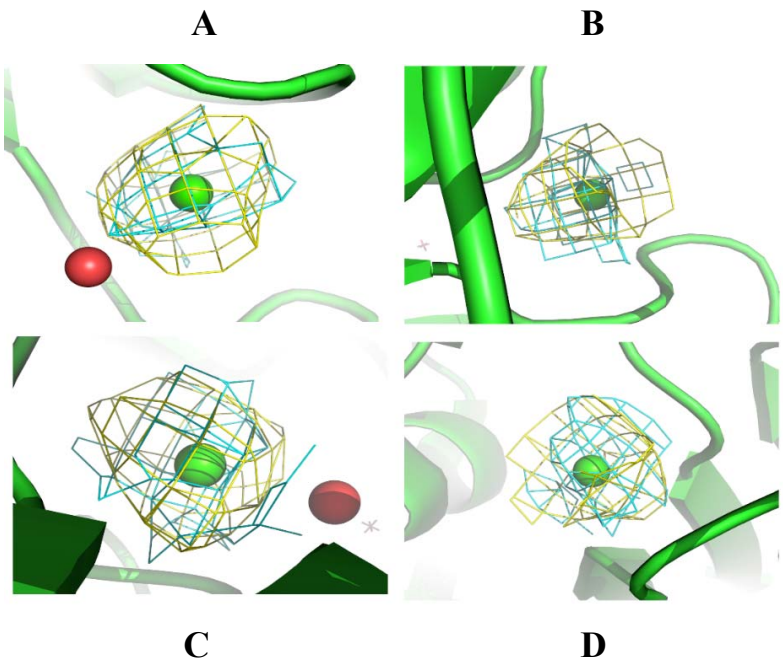

Figure S3 Electron density maps for the 4 Ca^{2+} ions. The figure displays electron density maps (2Fo-Fc map at 1σ level; Polder Fo-Fc omit map at 3σ level) for the 4 Ca^{2+} ions (A: Ca1501 in chain A; B: Ca1502 in chain A; C: Ca1501 in chain B; D: Ca1502 in chain B).

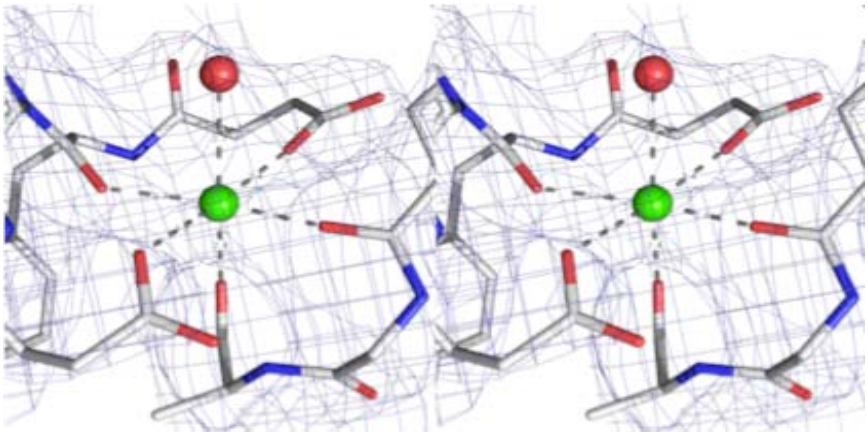

Figure S4 Representative electron density for Ca^{2+} coordination. This wall-eye stereo figure displays the coordination of the Ca^{2+} ion (Ca1501 in chain B) with representative electron density map (2Fo-Fc) contoured at 1σ level. The Ca^{2+} coordination is as follows (clockwise): 1. H_2O (top); 2. Asp1133 (side chain); 3. Lys1186 (main chain); 4. Ala1188 (main chain); 5. Glu1136 (main chain); 6. Tyr1134 (main chain). [See also Figure 2B and Table S1]

Table S1 Comparative analysis of calcium binding sites on C_{123}^{Pas}, $C_{123}^{AgI/II}$ and C_{123}^{SspB}.

Calcium	C_{123}^{Pas} (6E3F)			$C_{123}^{AgI/II}$ (3QE5)			C_{123}^{SspB} (7L0O)		
	Residue	**Atom**	**Distance (Å)**	**Residue**	**Atom**	**Distance (Å)**	**Residue**	**Atom**	**Distance (Å)**
at C_1/C_2	N910	OD1	2.5	N1155	OD1	2.5	ND	ND	ND
at C_1/C_2	Y911	O	2.5	Y1156	O	2.4	ND	ND	ND
at C_1/C_2	D944	OD1	2.8	D1189	OD1	2.7	ND	ND	ND
at C_1/C_2	D944	OD2	2.4	D1189	OD2	2.6	ND	ND	ND
at C_1/C_2	D946	OD1	2.7	D1191	OD1	2.7	ND	ND	ND
at C_1/C_2	Q947	OE1	2.7	Q1192	OE1	2.6	ND	ND	ND
in C_2	D967	OD1	2.7	D1212	OD2	2.6	D1133	OD1	2.4
in C_2	Y968	O	2.2	Y1213	O	2.6	Y1134	O	2.3
in C_2	E970	OE1	3.1	E1215	OE2	2.6	E1136	OE1	3.2
in C_2	K1020	O	2.4	K1265	O	2.8	K1186	O	2.4
in C_2	A1022	O	2.6	A1267	O	2.5	A1188	O	2.4
in C_3	ND	ND	ND	D1388	OD1	2.7	D1308	OD1	2.4
in C_3	ND	ND	ND	Y1389	O	2.3	Y1309	O	2.3
in C_3	ND	ND	ND	Q1391	OE1	2.8	ND	ND	ND
in C_3	ND	ND	ND	K1434	O	2.7	N1354	O	2.4
in C_3	ND	ND	ND	G1435	O	2.7	G1355	O	3.3

ND – Not Determined (These Ca^{2+} sites corresponding to Ca^{2+} sites in $C_{123}^{AgI/II}$ are not observed In C_{123}^{Pas} and C_{123}^{SspB}).

The table lists the conserved Ca^{2+} binding sites in the C_{123} domains of AgI/II, SspB and Pas.

Table S2 Analysis of results from protein-protein docking of C_{123}^{SspB} and $C_{123}^{AgI/II}$ to $SRCR_1$.

A. $C_{123}^{AgI/II}$ + $SRCR_1$ ($C_{12}^{AgI/II}$ interface): ΔG_{diss} = 3.0 kcal/mol

	Receptor $C_{123}^{AgI/II}$ (3QE5)	**Ligand $SRCR_1$ (6SA4)**
# of Interface residues	46	33
SASA interface area [Å^2]	1242 (5.9%)	1381 (26.0%)
Residues in the interface	K1023, Q1024, P1053, S1054, G1055, Y1056, Q1057, F1058, N1059, P1060, Y1074, N1076, A1077, N1079, G1110, Q1111, L1113, N1114, D1115, G1116, G1144, N1155, Y1156, I1157, K1158, T1160, D1189, D1191, L1246, S1258, K1259, A1260, G1261, I1262, R1263, P1264, K1265, Q1312, Y1314, N1320, G1321, Y1322, A1323, S1324, N1325, I1326	G99, L100, A101, L102, G109, R110, L118, Y119, R120, G121, Q142, L143, G144, R170, C171, S172, H174, E175, S176, Y177, W179, S180, C181, P182, H183, N184, G185, W186, A203, T204, G205, E206, N207

Hydrogen bonds	Receptor $C_{123}^{AgI/II}$ (3QE5)	Ligand $SRCR_1$ (6SA4)	Distance
1	LYS1158 [NZ]	TYR119 [OH]	2.58
2	ARG1263 [NE]	SER172 [OG]	3.55
3	GLN1024 [NE2]	TYR177 [OH]	2.72
4	TYR1314 [OH]	SER180 [O]	2.83
5	ASN1076 [ND2]	ALA203 [O]	3.52
6	ASN1076 [ND2]	GLU206 [OE2]	2.90
7	ASN1059 [ND2]	ASN207 [OD1]	2.74
8	GLN1057 [NE2]	ASN207 [O]	2.85
9	GLY1144 [O]	GLY99 [N]	3.25
10	ASP1115 [OD2]	LEU100 [N]	3.37
11	LYS1158 [O]	ARG120 [NH1]	2.81
12	ASP1189 [OD1]	ARG120 [NH1]	2.80
13	ASN1155 [OD1]	ARG120 [NH2]	2.70
14	TYR1156 [O]	ARG120 [NH2]	2.68
15	ASP1189 [OD2]	ARG120 [NH2]	2.70
16	SER1258 [O]	SER176 [N]	3.17
17	LYS1259 [O]	SER176 [OG]	2.76
18	SER1258 [O]	SER176 [OG]	3.80
19	ASN1320 [OD1]	TYR177 [OH]	2.71
20	GLY1321 [O]	TRP179 [NE1]	2.77
21	GLY1261 [O]	CYS181 [SG]	3.63
22	GLY1261 [O]	CYS181 [N]	3.02
23	PHE1058 [O]	ASN207 [N]	3.40
24	GLN1111 [OE1]	ASN207 [ND2]	2.88
Salt Bridges	**Receptor**	**Ligand**	**Distance**
1	ASP1189 [OD1]	ARG120 [NH1]	2.80
2	ASP1189 [OD2]	ARG120 [NH1]	2.90
3	ASP1189 [OD1]	ARG120 [NH2]	3.34
4	ASP1189 [OD2]	ARG120 [NH2]	2.70

B. C_{123}^{SspB} + $SRCR_1$ (C_{12}^{SspB} interface): ΔG_{diss} = 4.3 kcal/mol

	Receptor C_{123}^{SspB} (7L0O)	**Ligand $SRCR_1$ (6SA4)**
# of Interface residues	34	27
SASA interface area [Å^2]	983 (4.6%)	1038 (20.1%)
Residues in the interface	K944, L1034, N1035, D1036, G1037, G1065, Y1077, I1078, K1079, P1080, T1081, D1093, G1094, D1133, Y1134, P1135, E1136, E1137, A1181, N1182, I1183, T1184, K1186, E1231, K1233, Y1235, G1242, Y1243, A1244, T1245, E1246, V1247, V1248, V1249	G99, L100, L118, Y119, R120, G121, N134, R141, D168, R170, C171, S172, G173, H174, E175, S176, Y177, W179, S180, C181, P182, H183, N184, G185, W186, L187, S188

Hydrogen bonds	**Receptor** C_{123}^{SspB} **(7L0O)**	**Ligand** $SRCR_1$ **(6SA4)**	**Distance**
1	LYS1233 [NZ]	TYR119 [OH]	2.50
2	TYR1235 [OH]	TYR119 [OH]	3.66
3	ASN1035 [ND2]	ASN134 [OD1]	3.87
4	ASN1182 [ND2]	ASP168 [OD1]	3.02
5	ASN1035 [ND2]	SER172 [O]	2.95
6	ALA1244 [N]	SER180 [O]	3.27
7	ASN1182 [ND2]	HIS183 [O]	3.14
8	THR1184 [OG1]	ASN184 [O]	2.85
9	LYS1186 [NZ]	TRP186 [O]	2.60
10	LYS1186 [NZ]	LEU187 [O]	2.53
11	VAL1247 [O]	GLY99 [N]	2.88
12	VAL1247 [O]	LEU100 [N]	2.99
13	ASP1133 [OD2]	TYR119 [OH]	2.80
14	TYR1156 [O]	ARG120 [N]	2.83
15	ASP1133 [OD1]	ARG120 [NH1]	2.83
16	LYS1186 [O]	ARG120 [NH1]	3.70
17	GLU1231 [O]	ARG120 [NH2]	2.69
18	ASP1133 [O]	ARG120 [NH2]	2.64
19	GLU1231 [OE2]	GLY121 [N]	2.99
20	TYR1077 [OH]	ARG141 [NE]	2.87
21	ASP1036 [OD2]	HIS174 [ND1]	3.49
22	ILE1078 [O]	HIS174 [NE2]	3.40
23	GLU1246 [OE1]	TYR177 [OH]	2.86
24	ALA1244 [O]	SER180 [O]	2.84
25	ASN1182 [O]	HIS183 [ND1]	2.81
Salt Bridges	**Receptor**	**Ligand**	**Distance**
1	ASP1133 [OD1]	ARG120 [NH1]	2.83
2	ASP1133 [OD2]	ARG120 [NH1]	2.89
3	ASP1133 [OD1]	ARG120 [NH2]	3.83
4	ASP1133 [OD2]	ARG120 [NH2]	2.80
5	ASP1036 [OD2]	HIS 174 {ND1]	3.49

C. C_{123}^{SspB} + $SRCR_1$ (C_{23}^{SspB} interface): ΔG_{diss} = 4.9 kcal/mol

	Receptor C_{123}^{SspB} (7L0O)	**Ligand $SRCR_1$ (6SA4)**
# of Interface residues	43	28
SASA interface area [Å^2]	1099 (5.2%)	1208 (23.5%)
Residues in the interface	K1087, A1099, G1100, S1101, T1102, Y1104, D1140, R1142, D1144, L1145, T1218, V1219, K1220, S1221, E1222, K1225, K1259, E1301, D1302, S1304, V1306, D1308, Y1309, D1310, Q1311, D1351, T1353, N1354, I1356, T1358, R1360, E1392, N1393, T1394, Y1395, V1396, T1398, N1400, K1401, V1402, A1403, A1405, S1406	G99, L100, A101, L102, G109, R110, L118, Y119, R120, G121, R141, Q142, L143, G144, H174, S176, Y177, W179, S180, C181, P182, H183, W186, A203, T204, G205, E206, N207

Hydrogen bonds	Receptor C_{123}^{SspB} (7L0O)	Ligand $SRCR_1$ (6SA4)	Distance
1	ARG1360 [NH1]	SER176 [OG]	2.87
2	ARG1360 [NH2]	SER176 [OG]	2.73
3	SER1304 [OG]	TYR177 [OH]	2.62
4	THR1398 [OG1]	TYR177 [OH]	2.75
5	ASN1354 [ND2]	TRP179 [O]	2.76
6	ASN1354 [ND2]	CYS181 [O]	2.89
7	ARG1142 [NH2]	ALA203 [O]	2.71
8	ARG1142 [NH2]	GLY205 [O]	2.79
9	LYS1087 [NZ]	GLU206 [OE1]	2.46
10	TYR1104 [OH]	GLU206 [OE2]	2.69
11	ARG1142 [NH1]	GLU206 [OE2]	2.82
12	THR1102 [N]	GLU206 [O]	3.26
13	THR1102 [OG1]	GLU206 [O]	2.97
14	SER1406 [O]	GLY99 [N]	3.01
15	GLY1100 [O]	ARG110 [NH1]	3.07
16	GLU1392 [OE1]	TYR119 [OH]	2.85
17	ASP1308 [OD1]	ARG120 [NE]	2.93
18	GLU1392 [O]	ARG120 [NH1]	3.53
19	ASP1310 [OD2]	ARG120 [NH1]	2.92
20	GLU1392 [OE2]	ARG120 [NH1]	2.80
21	ASP1308 [OD2]	ARG120 [NH2]	2.77
22	TYR1309 [O]	ARG120 [NH2]	2.97
23	ASP1308 [O]	ARG120 [NH2]	3.25
24	LYS1401 [O]	GLY144 [N]	2.71
25	ASN1354 [OD1]	HIS183 [N]	3.01
26	ASP1308 [OD2]	TRP186 [NE1]	3.88
27	GLU1301[OE2]	ASN207 [ND2]	2.90
Salt Bridges	**Receptor**	**Ligand**	**Distance**
1	LYS1087 [NZ]	GLU206 [OE1]	2.46
2	LYS1087 [NZ]	GLU206 [OE2]	2.72
3	ARG1142 [NH1]	GLU206 [OE2]	2.82
4	ASP1308 [OD2]	ARG120 [NE]	3.96

Table S3 Analysis of results from protein-protein docking of C_{123}^{SspB} to P. gingivalis shaft protein Mfa1, and comparison of binding region in C_{123}^{SspB} with the corresponding region in $C_{123}^{AgI/II}$ and C_{123}^{Pas}.

	Receptor Mfa1 (5NF2)	**Ligand C_{123}^{SspB} (7L0O)**
# of Interface residues	34	30
SASA interface area [Å²]	916 (4.6%)	928 (4.1%)
Residues in the interface	K246, A247, S249, Y250, E251, K253, A254, T255, A263, G264, V266, T343, L344, A345, Q348, T470, T471, S472, P473, D474, S475, W476, H490, K492, S493, K495, Q541, D542, T543, M545, S546, V547, E548, V549	K1121, E1122, Q1125, N1126, G1127, Y1129, P1143, D1144, V1146, Q1155, S1157, G1158, V1159, S1160, V1161, Q1162, Y1164, A1170, P1171, **K1172, K1173, V1174,** L1191, S1193, D1195, N1196, P1197, E1198, E1199, K1202

	#	**Receptor Mfa1 (5NF2)**	**Ligand C_{123}^{SspB} (7L0O)**	**Distance**
Hydrogen bonds	1	LYS246 [NZ]	SER1193 [OG]	2.33
	2	SER493 [OG]	ASP1195 [OD2]	2.84
	3	THR543 [N]	GLU1199 [OE1]	3.76
	4	THR543 [OG1]	GLU1199 [OE1]	3.41
	5	ASP474 [OD1]	LYS1202 [NZ]	3.64
Salt bridge	1	ASP474 [OD1]	LYS1202 [NZ]	3.64

Sequence alignment of 3QE5 (AgI/II), 6E3F (Pas) and 7L0O (SspB).

The boxed area highlights the BAR peptide motif, where high homology exists between AgI/II and Pas, but which is very different SspB.

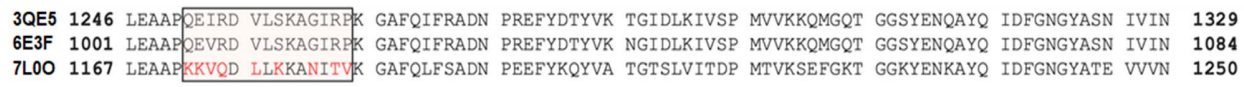

Table S4 Residue range for the structures of C_{23}^{SspB} and C_{123}^{SspB}.

PDBcode	Domains	Residues	Isopeptide bond in C_1	Isopeptide bond in C_2	Isopeptide bond in C_3
2WZA	C_2-C_3	1076-1413	-N/A-	K1082-1232	K1259 - N1393
2WOY	C_2-C_3	1081-1413	-N/A-	-None-	K1259 - N1393
This study	C_1-C_3	914 - 1406	K927 - N1042	K1082 - N1232	-None-

Table S5 Equivalent protein residues involved in glucose binding.

	$C_{123}^{AgI/II}$ (3QE5)	C_{123}^{SspB} (7L0O)
GLC1	K1023	K944
	Q1024	Q945
	L1113	L1034
	K1259	K1180
	A1260	A1181
	G1261	N1182
	G1319	G1240
	N1320	N1241
	G1321	G1242
GLC2	K1023	K944
	D1115	D1036
	Y1156	Y1077
	I1157	I1078
	K1158	K1079
	G1321	G1242
	Y1322	Y1243
	A1323	A1244
GLC3	S1054	T975
	G1055	G976
	Y1056	Y977
	Q1057	K978
	F1058	F979
	Y1074	Y995
	N1076	E997
	N1079	H1000
	L1113	L1034

GLC4	Q1003	Q924
	K1134	K1055
	S1135	S1056
	N1136	N1057
	V1137	V1058
GLC5	R1018	R939
	R1139	R1060
	P1149	P1070
	D1150	D1071

The table lists residues in C_{123}^{SspB} (7L0O) and C_{123}^{Pas} (6E3F), which are equivalent to residues in the structure of $C_{123}^{AgI/II}$ (3QE5) that interact with glucose molecules. These interaction sites may point to a potential region for recognition of the highly glycosylated Gp340.