1 SUPPLEMENTARY MATERIAL

The rvET optimized alignments when input into ivET or Shannon Entropy ranking method increased the average z-score $\langle z_o \rangle$ of functional overlap from 2.98 to 3.45 (16%) and from 3.61 to 3.82 (6%), respectively (see Figure 1). The ivET and Shannon Entropy optimized sequence selections also triggered improvements in rvET, but these were quite slight (1-4%). This may be explained by the intrinsic robustness of rvET compared to the coarseness of ivET and Shannon Entropy methods.
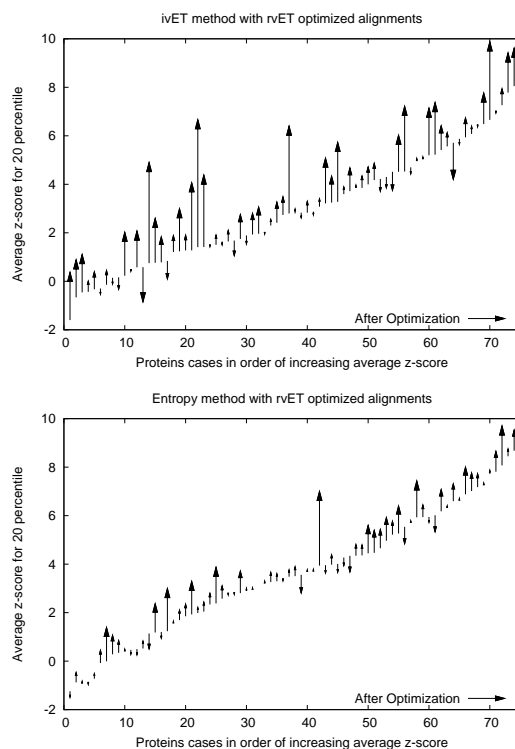

Figure 1: The ivET and Shannon Entropy ranking methods were tested on the rvET optimized alignments. The sequence selection optimization using the rvET also helps the control methods improve site prediction.

1

Testset 1		Number of sequences			
PDBID	unoptimized	optimized	PDBID	unoptimized	optimized
16pk	492	132	1dam	462	119
1a09	295	179	1dig	467	207
1a0oE	481	136	1dqr	467	169
1a22A	319	137	1dqx	152	81
1a22B	141	49	1e96A	439	144
1a2kA	124	77	1e96B	62	41
1a2kD	424	218	1ee9	164	109
1a3k	349	127	1efaB	470	73
1a48	425	121	1eg2	264	85
1a4mA	295	89	1eje	126	81
1a53	468	145	1elrA	385	182
1a59	478	216	1elwA	406	137
1a6m	344	212	1f6mA	488	180
1a6q	238	162	1f88A	260	118
1a80	482	199	1finA	444	163
1aca	352	219	1finB	417	204
1ad3A	459	224	1fjmA	422	167
1ai2	350	192	1fqjB	281	133
1aj2	471	143	1gnjA	92	41
1aj8A	475	181	1jfiB	133	74
1aky	454	226	1k7vA	206	87
1am1	354	156	1ng1	464	204
1amk	416	148	1nzcA	478	120
1aonF	475	261	1pvdA	239	127
1ars	388	162	1qumA	310	133
1aru	93	49	1qupA	47	34
1ast	366	125	1rrpA	411	196
1axn	441	157	1rrpB	207	99
1b54	483	94	1vh4A	243	116
1bag	41	24	1w1uA	319	170
1bqk	63	42	1ycsA	112	60
1bto	490	184	1ycsB	57	49
1c1bA	474	301	2bif	253	96
1cg0	479	244	2mjpA	488	325
1cio	396	220	2msbA	298	142
1cvjA	310	139	3hhrA	339	173
1cxzA	446	176	6gst	361	113

Table 1: The change in sequence count for training set due to the optimization is shown.

Testset 2		Number of sequences			
PDBID	unoptimized	optimized	PDBID	unoptimized	optimized
1aa6	438	193	1fca	428	138
1aac	126	89	1ffh	465	188
1ah7	29	24	1fit	170	98
1ako	480	166	1fnc	200	124
1amj	255	135	1fsu	182	69
1apq	254	127	1fxd	24	19
1arv	89	56	1gai	79	61
1at0	82	64	1gcb	151	68
1ayl	417	179	1gpl	192	100
1bdb	482	196	1han	153	91
1bia	225	114	1htn	219	111
1bif	240	118	1hyt	323	96
1bip	64	41	1iba	130	67
1bor	26	19	1ido	280	110
1btl	449	201	1ig5	51	28
1cfb	183	96	1iyu	335	127
1chc	93	49	1krn	190	92
1chd	457	164	1lam	406	240
1csh	200	119	1lay	31	23
1ctn	145	58	1lcf	191	93
1ctt	79	51	1led	349	123
1cvl	116	67	1lgr	455	180
1def	474	180	1lml	178	74
1drw	469	225	1mla	487	177
1dxy	449	176	1mup	86	41
1e70	449	186	1nif	37	31
1ecl	473	171	1nir	34	27
1emn	1277	36	1nox	49	27
1esl	73	39	1onc	43	25
1far	81	32	1opc	422	125

Table 2:   The change in sequence count for testset due to the optimization is shown.

3

Testset 2		Number of sequences			
PDBID	unoptimized	optimized	PDBID	unoptimized	optimized
1osa	397	204	1vii	47	40
1pbn	331	144	1vsd	190	56
1pda	458	211	1whi	493	188
1pdc	218	147	1xnb	243	115
1pii	156	84	2abk	488	203
1pkp	465	233	2ace	409	139
1poa	379	149	2af8	46	35
1poc	38	30	2asi	306	96
1pth	83	53	2cba	376	151
1put	443	124	2cmd	281	147
1qli	326	115	2dkb	391	209
1rfs	72	56	2dln	447	146
1rie	395	179	2fha	399	165
1rnl	487	262	2hft	32	25
1se4	87	44	2rn2	479	125
1snc	164	61	2sil	48	23
1sp2	164	128	2vil	286	215
1sra	63	44	3dni	110	58
1thg	285	106	3ebx	176	54
1thm	441	358	3ssi	31	26
1thx	483	271	4enl	446	179
1tmy	460	305	4rhn	474	205
1uch	165	62	5eat	409	147
1uxc	130	82	5ptp	417	139
1vhh	66	33	7rsa	341	114

Table 3: The change in sequence count for testset due to the optimization is shown.
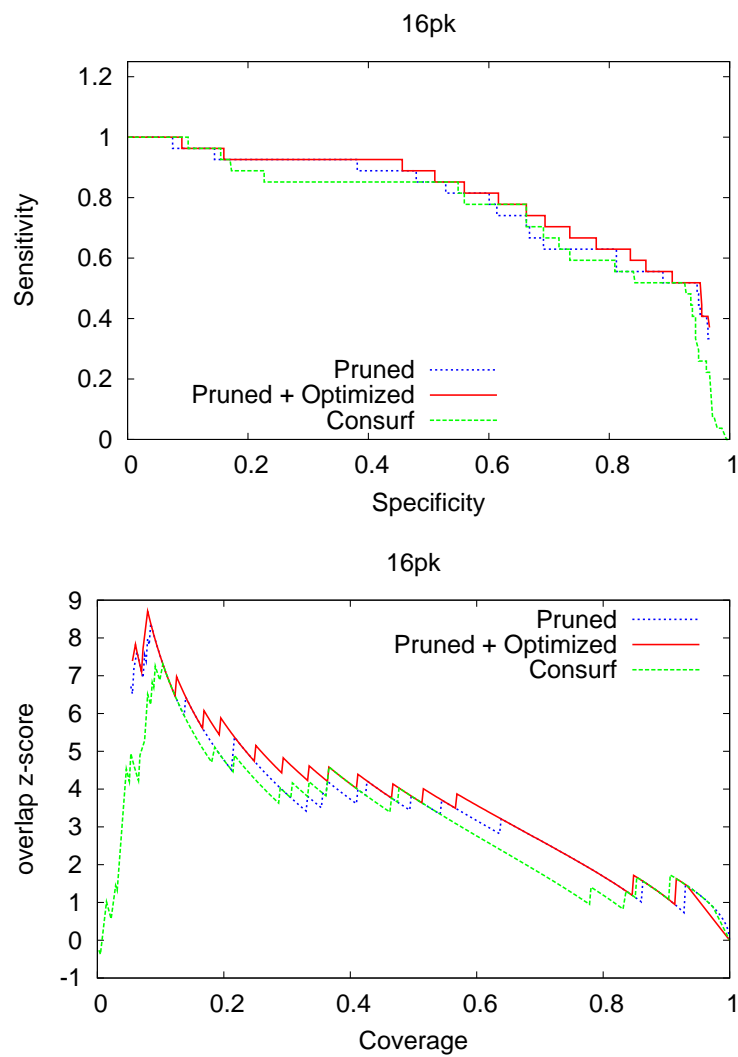
Figure 2: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 16pk. b) The overlap z-scores as a function of coverage for the three methods is shown.
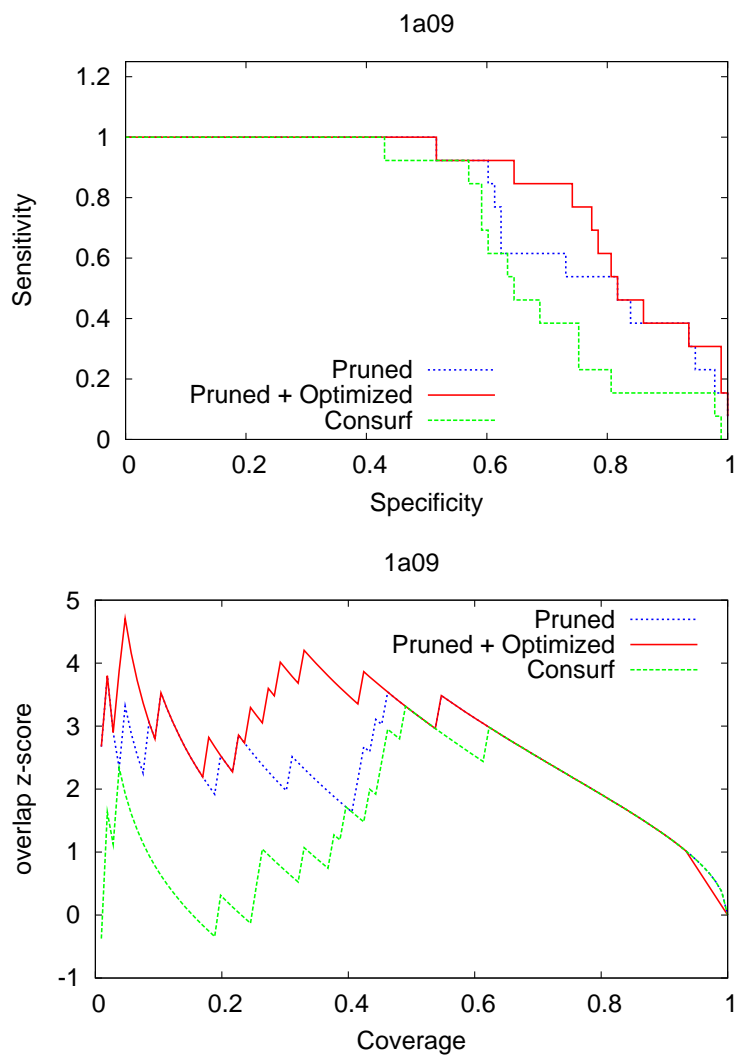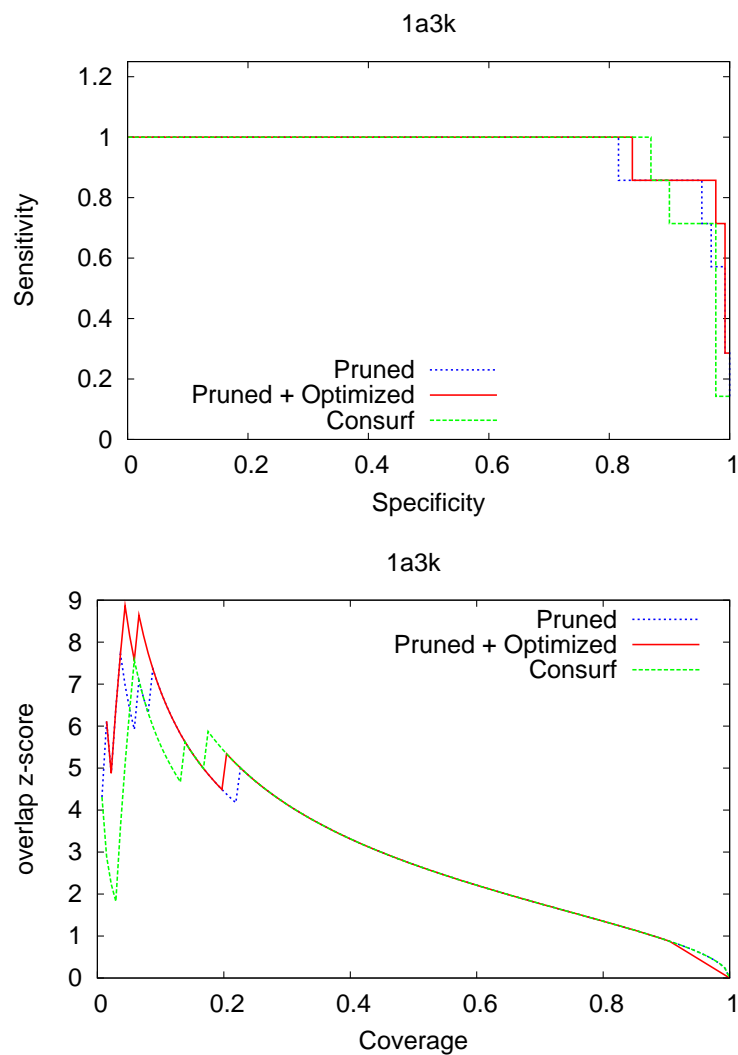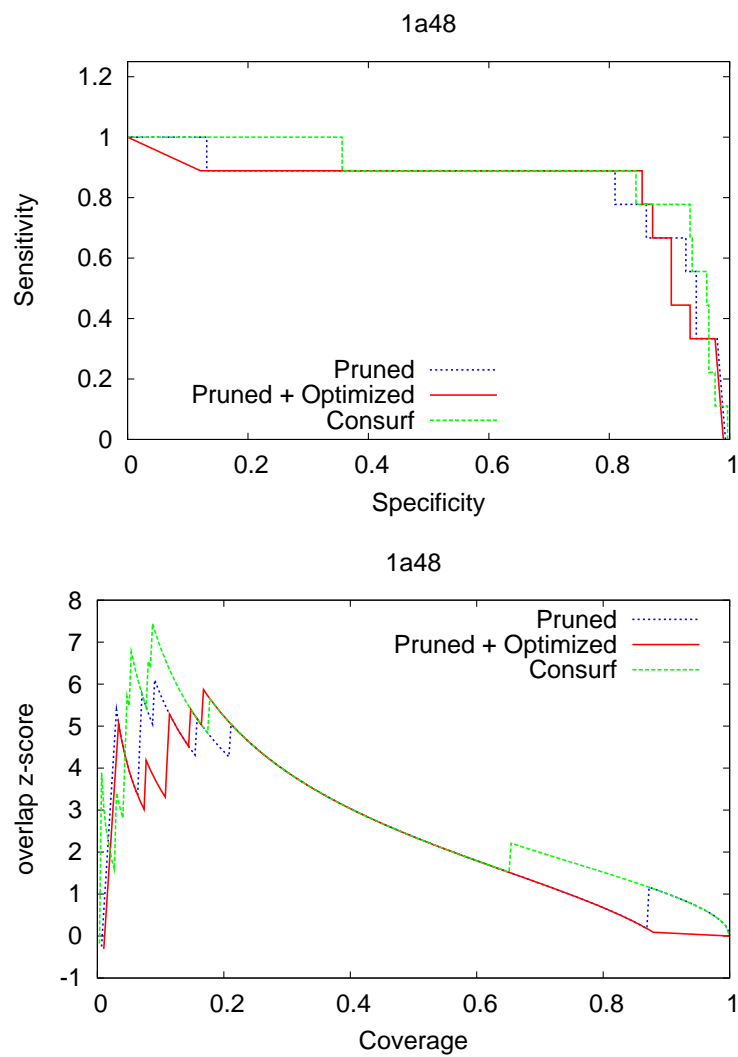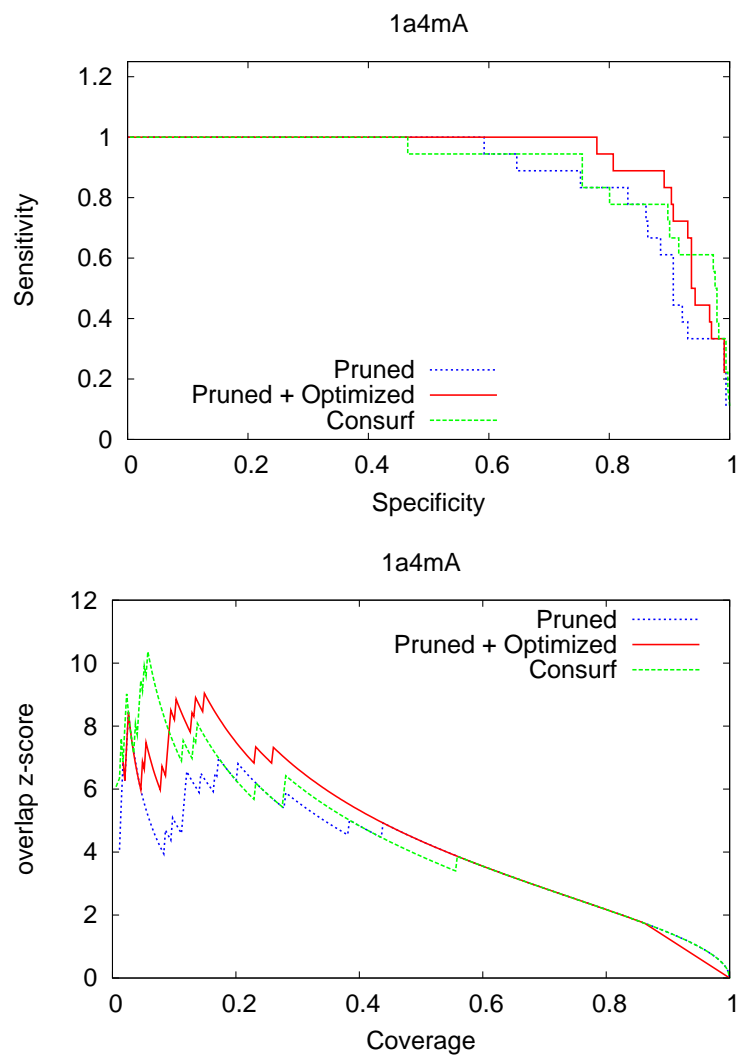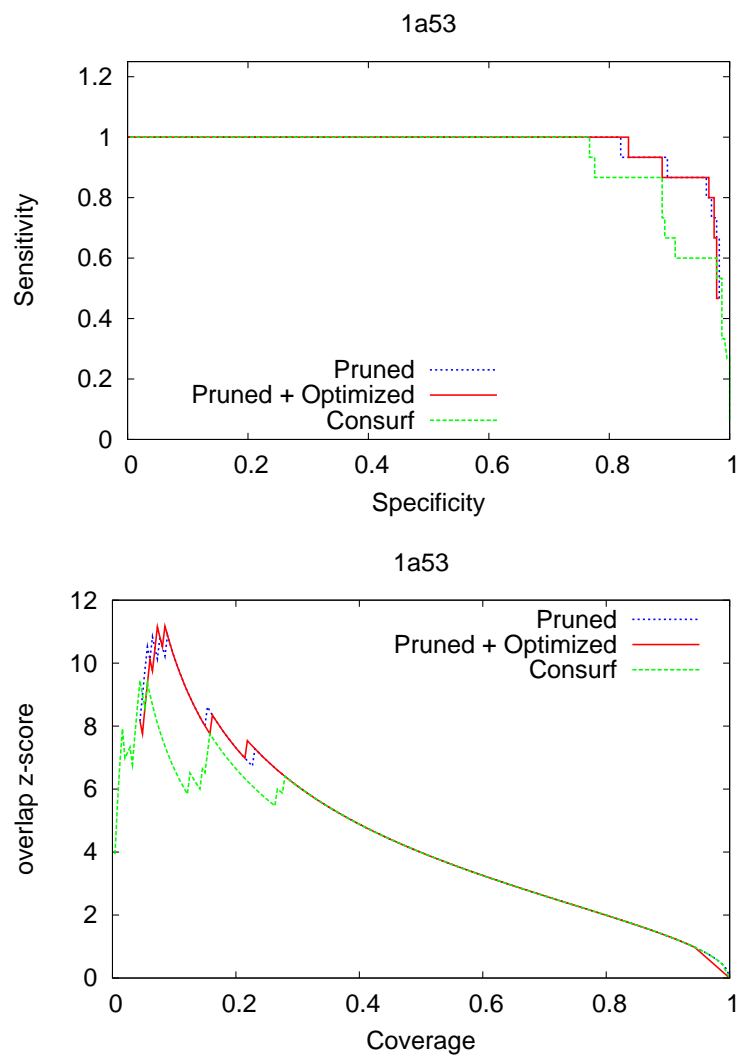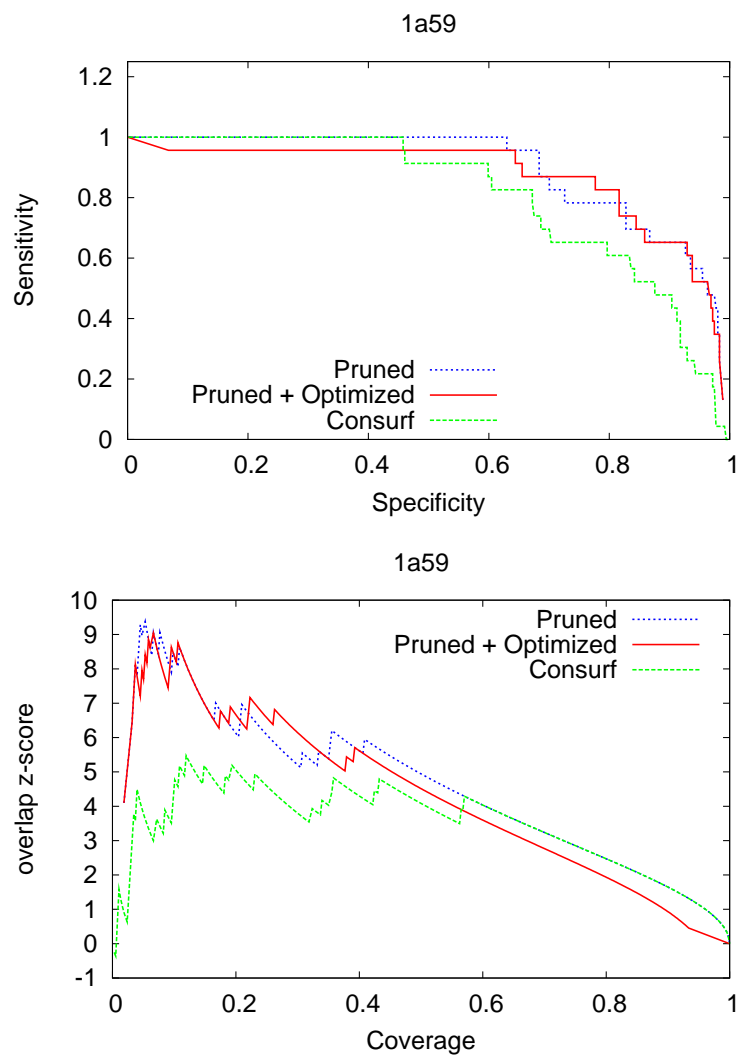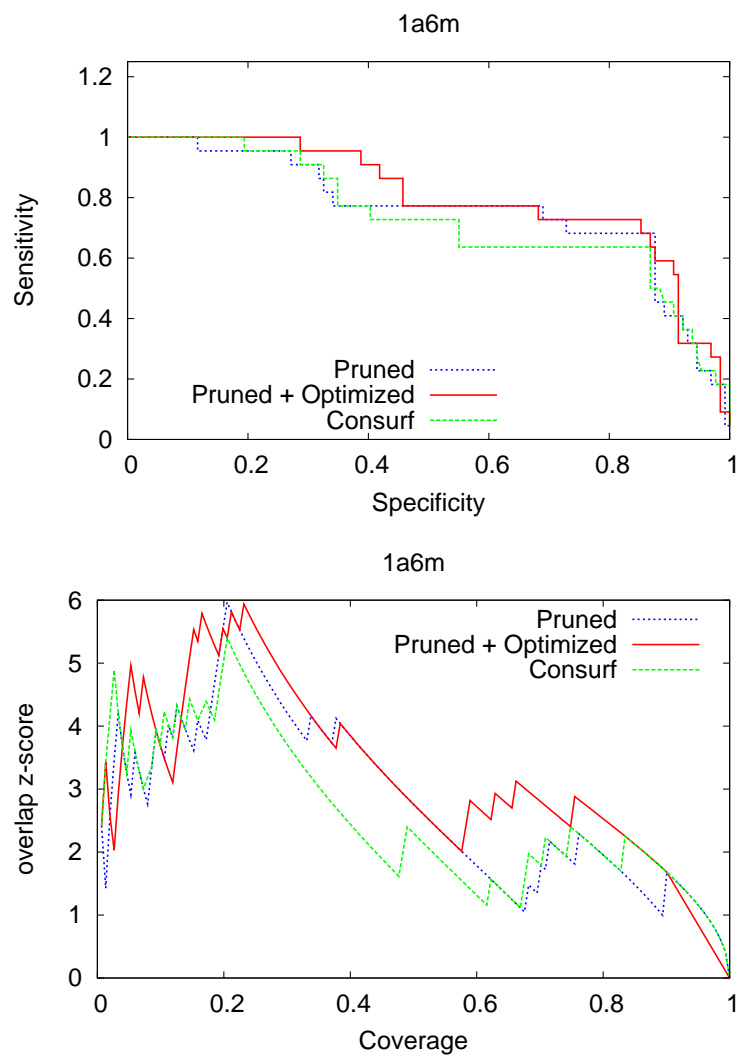
Figure 3: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a09. b) The overlap z-scores as a function of coverage for the three methods is shown.
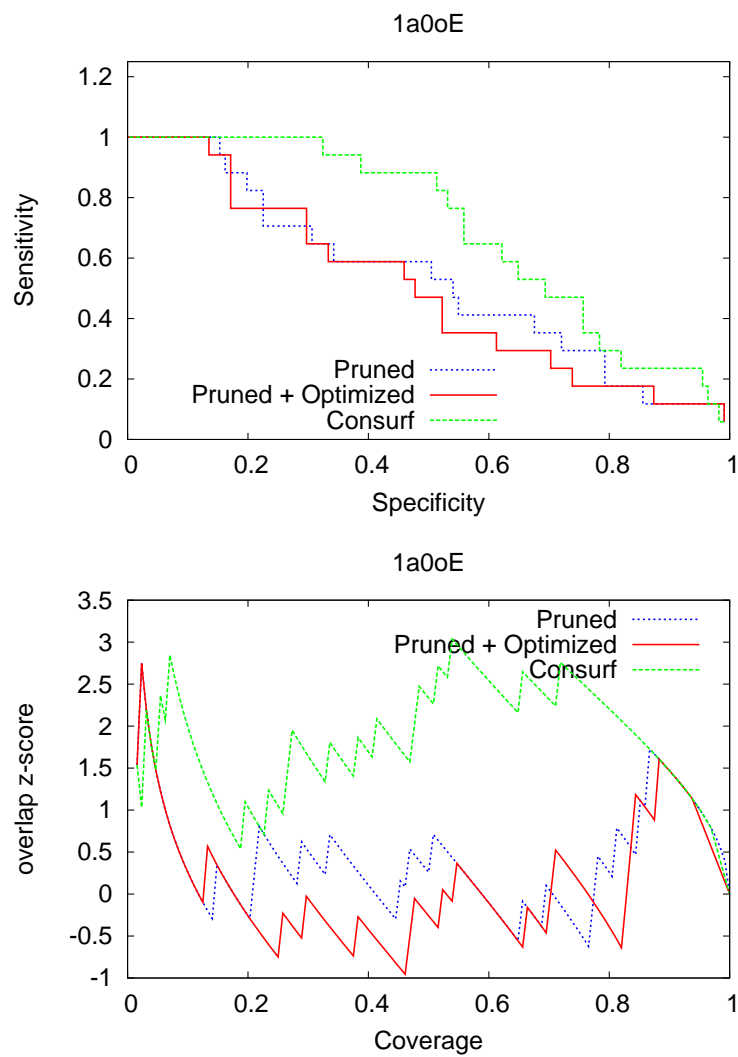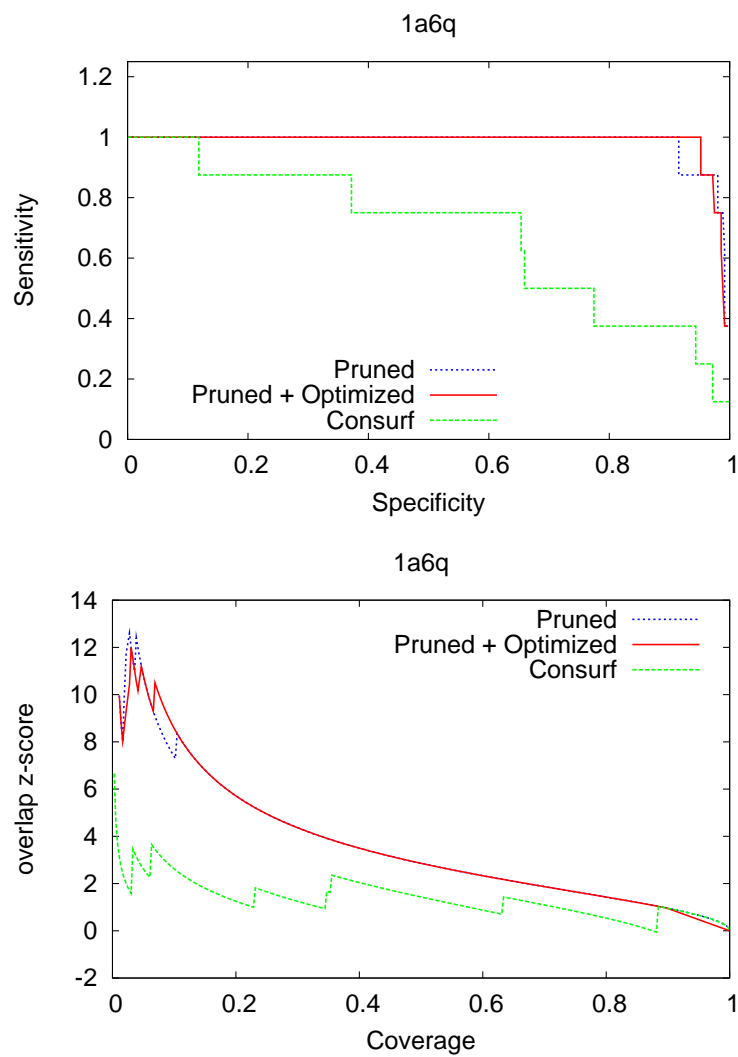
Figure 4: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a0oE. b) The overlap z-scores as a function of coverage for the three methods is shown.
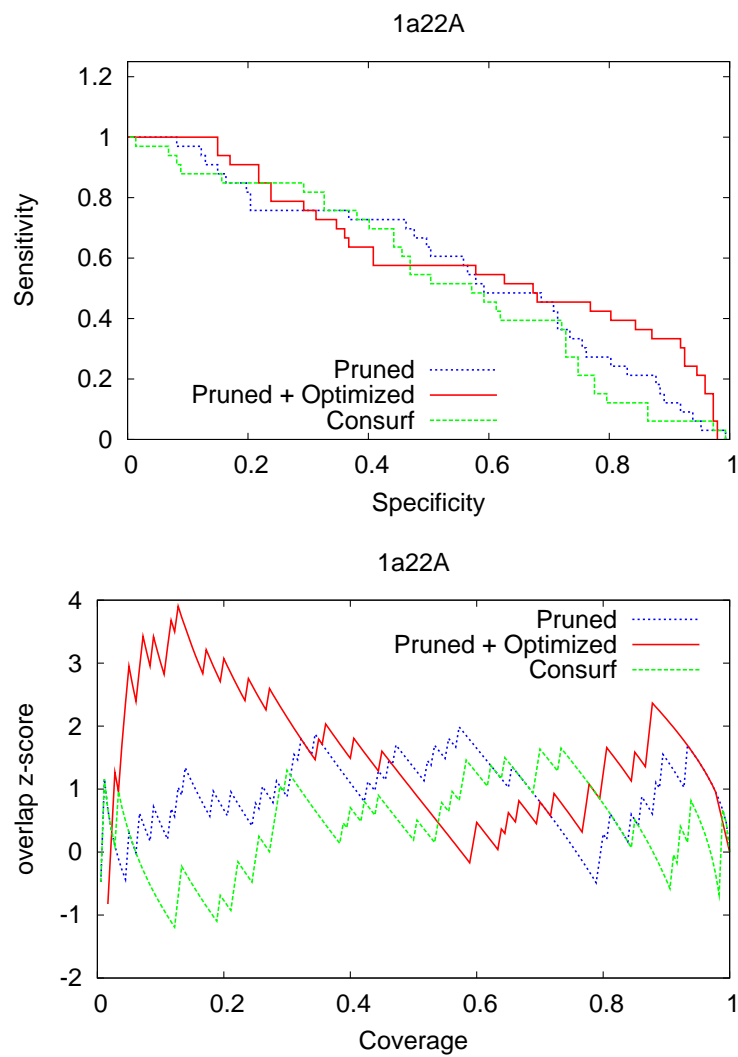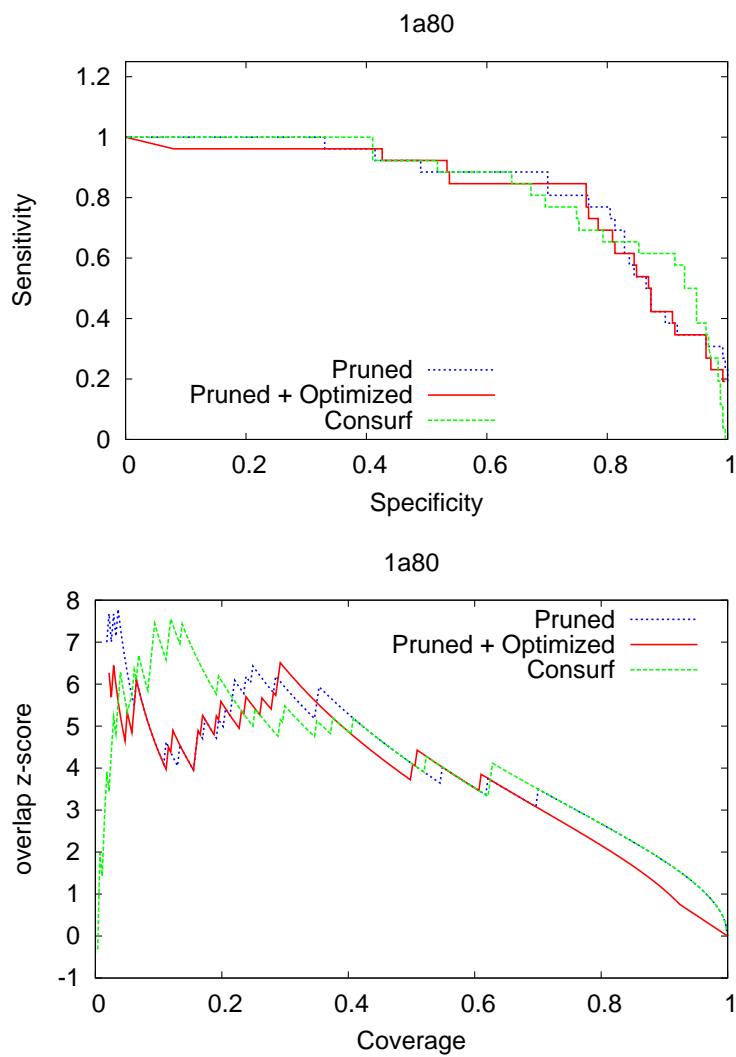
Figure 5: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a22A. b) The overlap z-scores as a function of coverage for the three methods is shown.
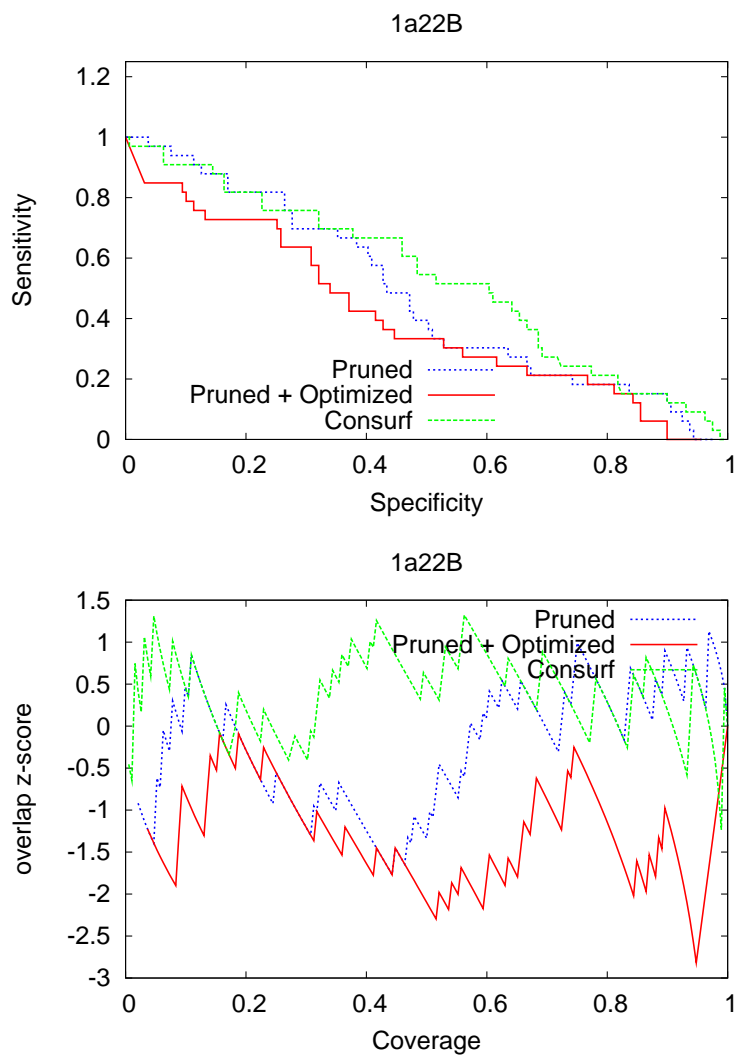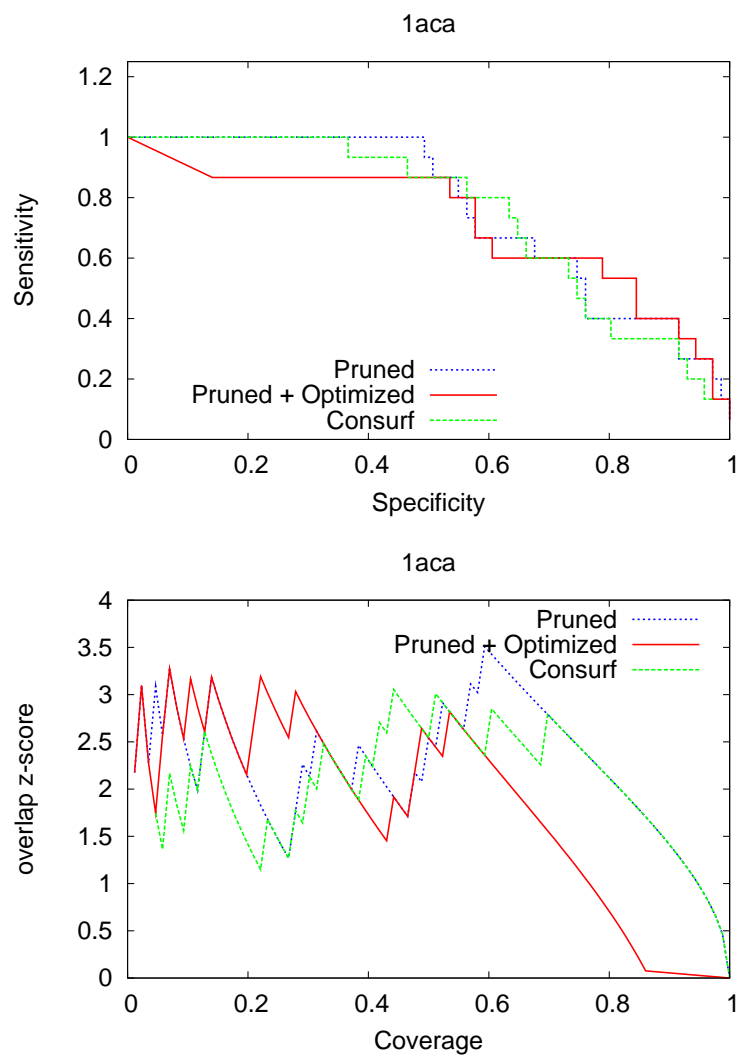
Figure 6: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a22B. b) The overlap z-scores as a function of coverage for the three methods is shown.
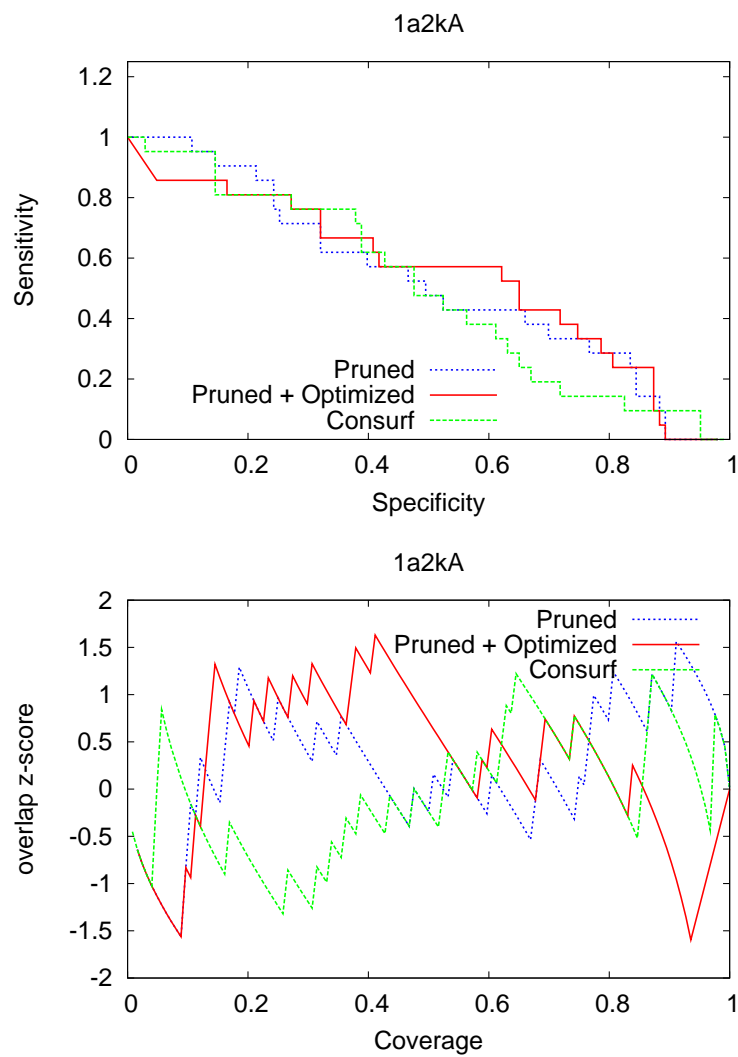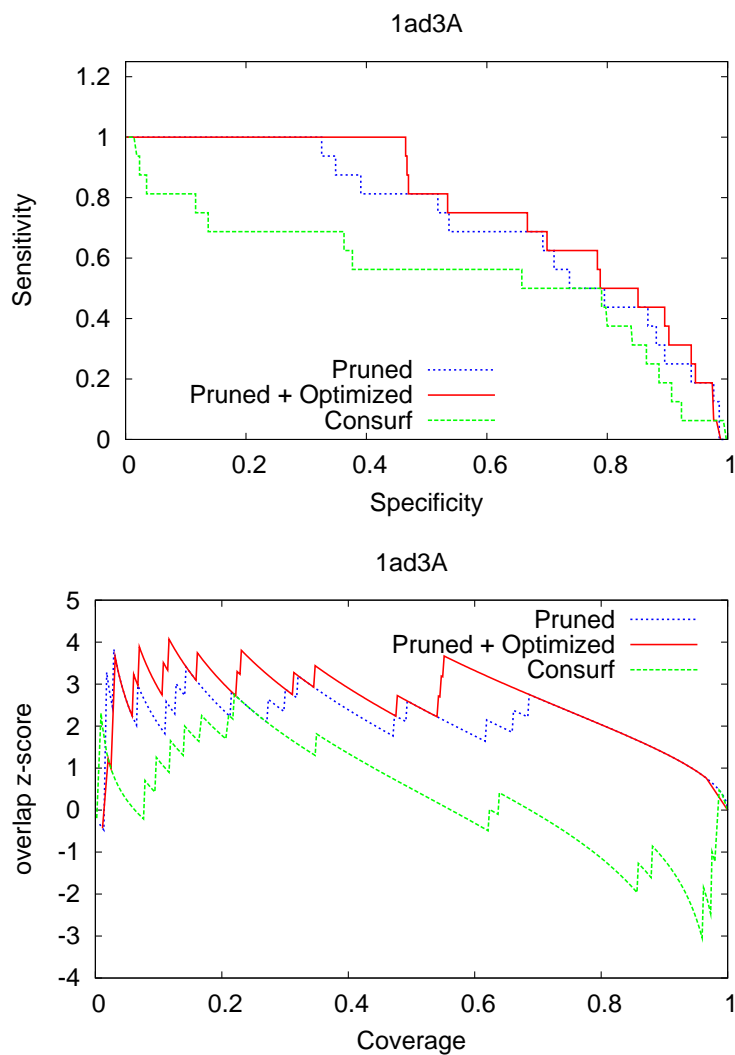
Figure 7: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a2kA. b) The overlap z-scores as a function of coverage for the three methods is shown.
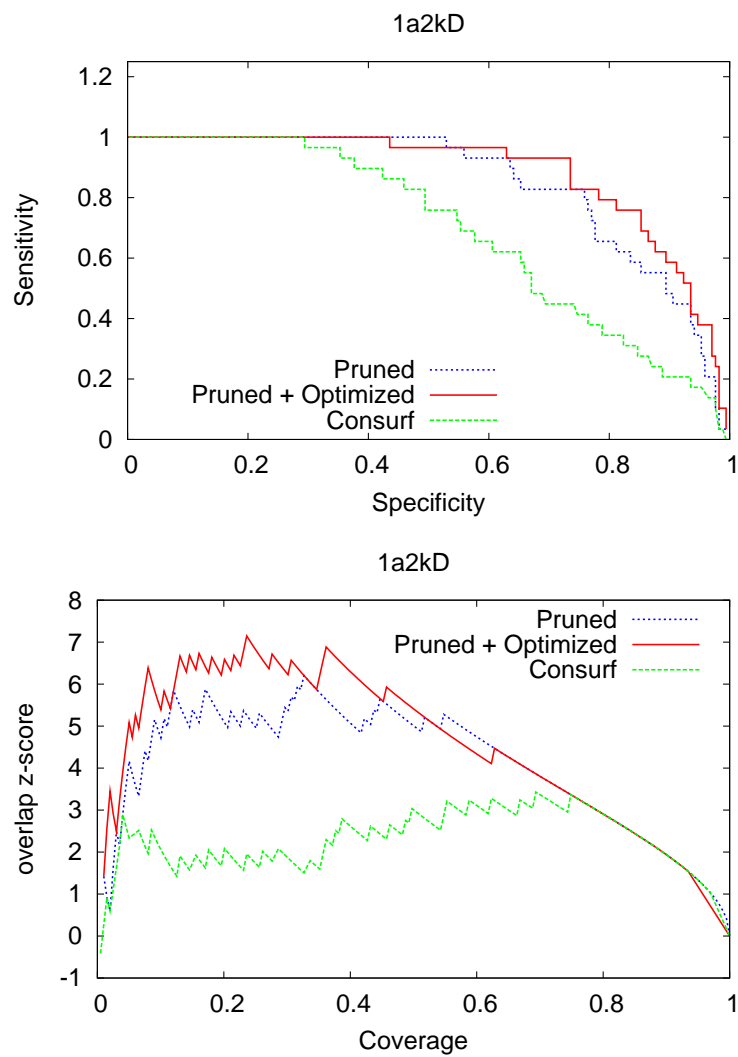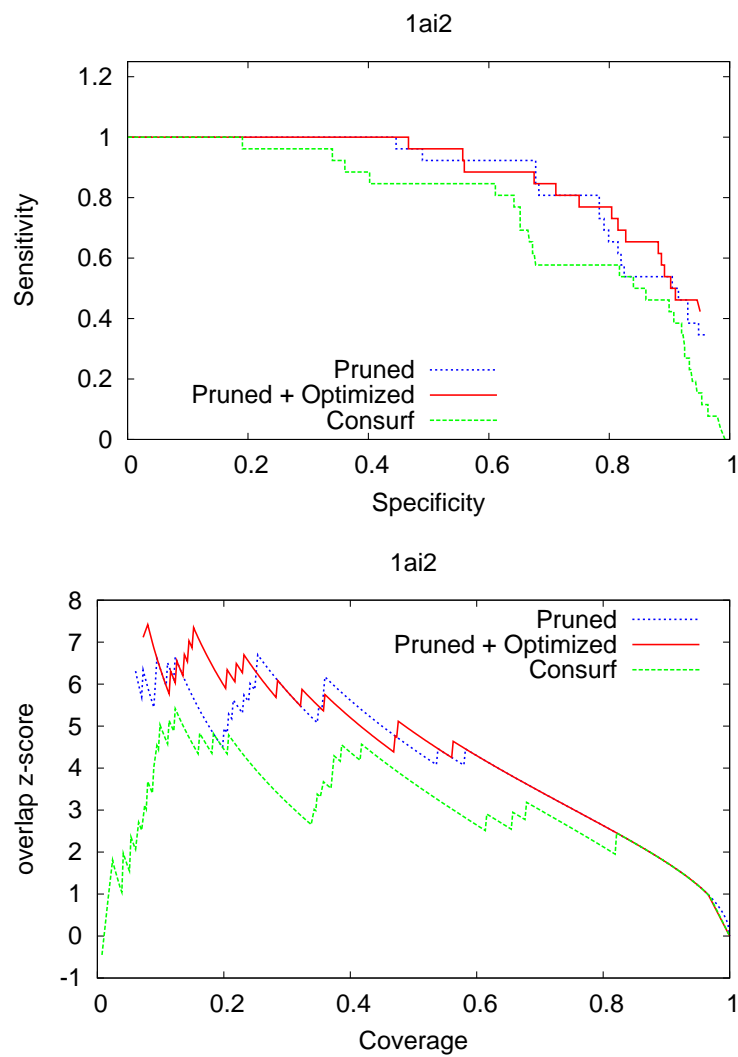
Figure 8: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a2kD. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 9: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a3k. b) The overlap z-scores as a function of coverage for the three methods is shown.
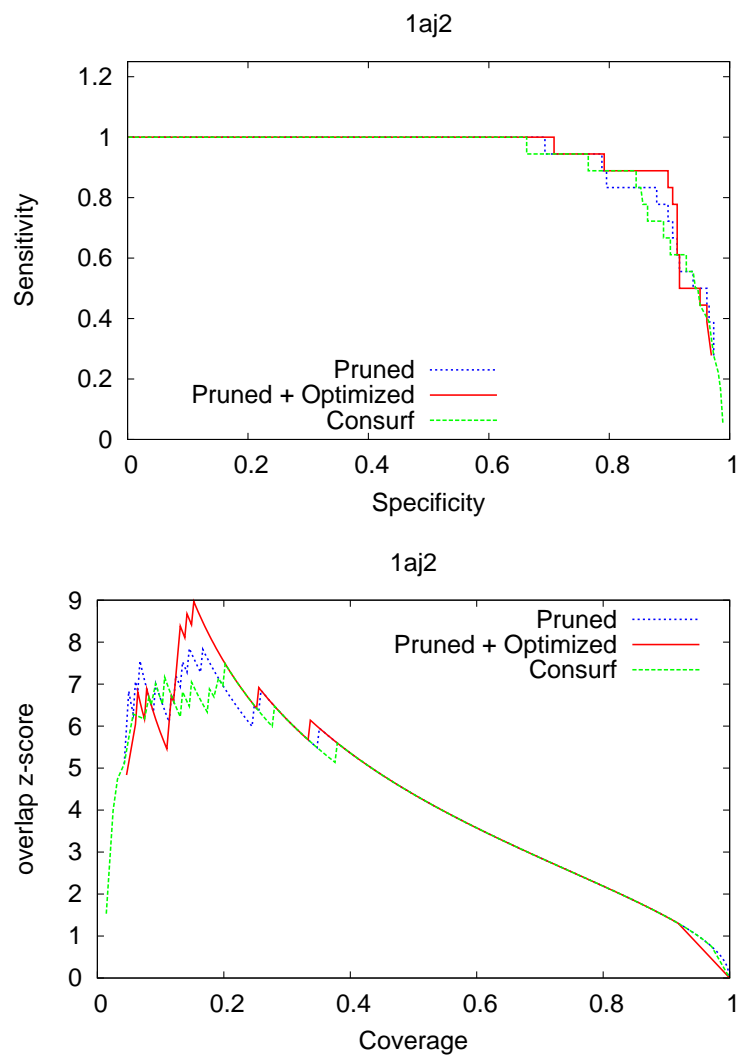
Figure 10: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a48. b) The overlap z-scores as a function of coverage for the three methods is shown.
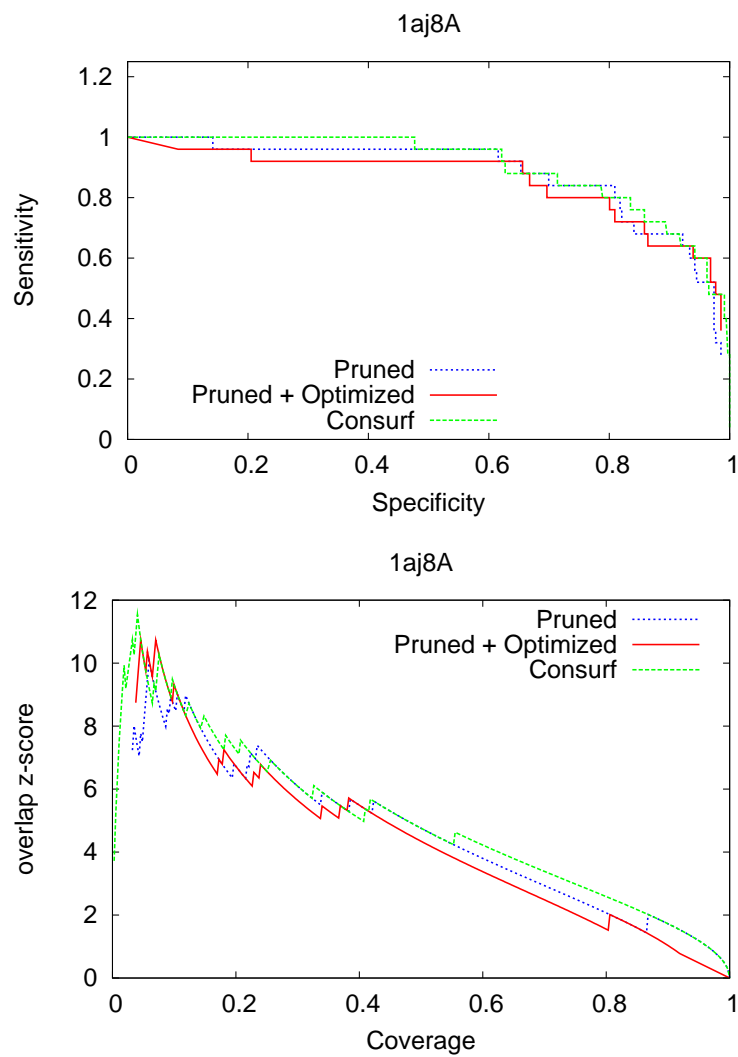
Figure 11: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a4mA. b) The overlap z-scores as a function of coverage for the three methods is shown.
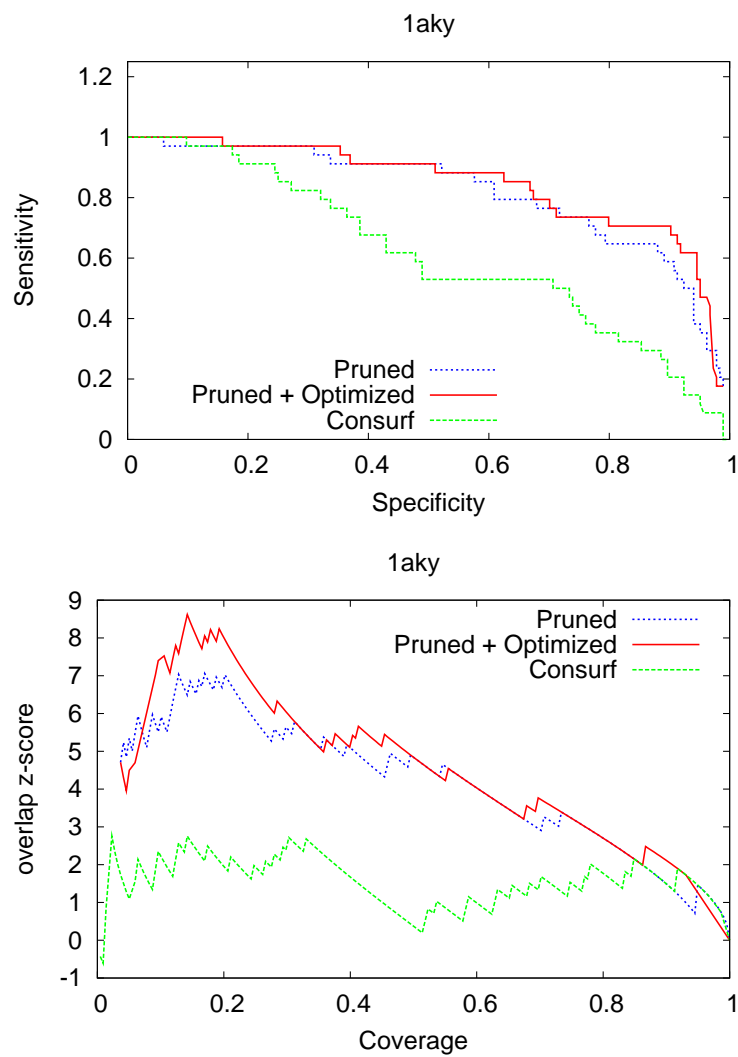
Figure 12: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a53. b) The overlap z-scores as a function of coverage for the three methods is shown.
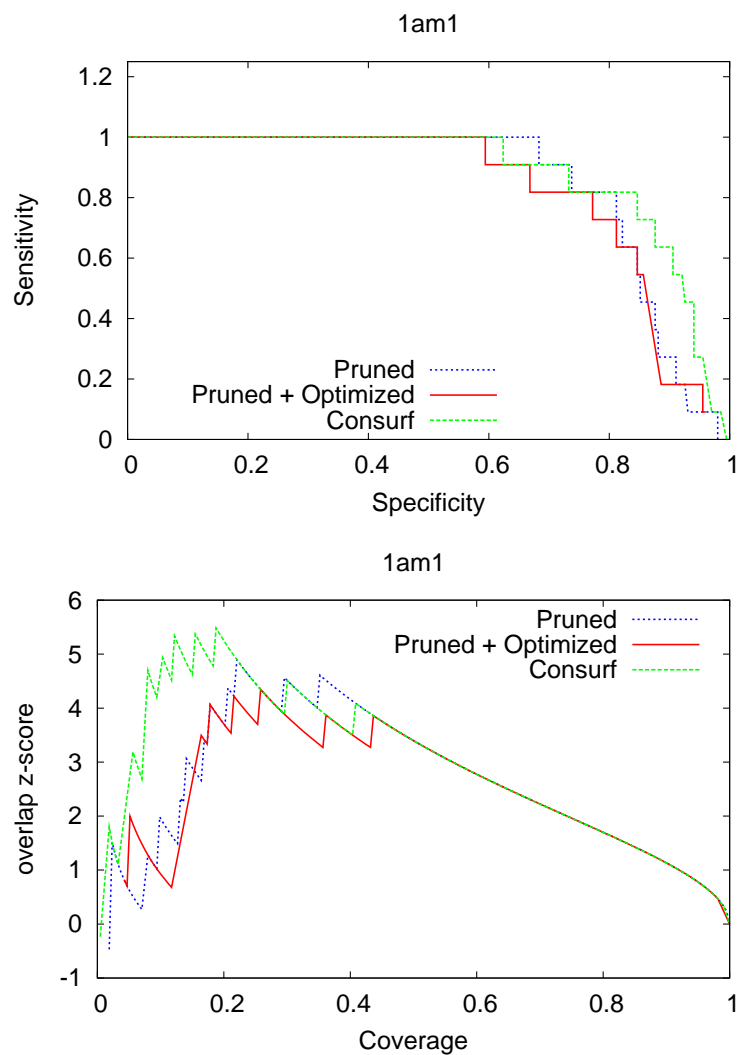
Figure 13: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a59. b) The overlap z-scores as a function of coverage for the three methods is shown.
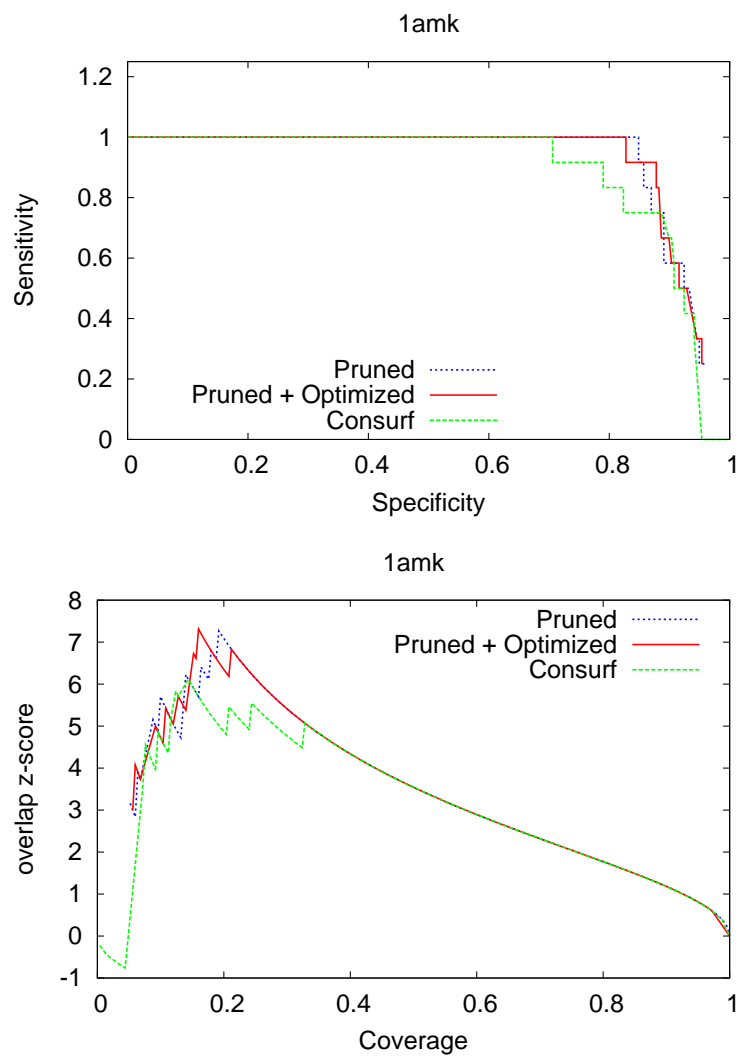
Figure 14: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a6m. b) The overlap z-scores as a function of coverage for the three methods is shown.
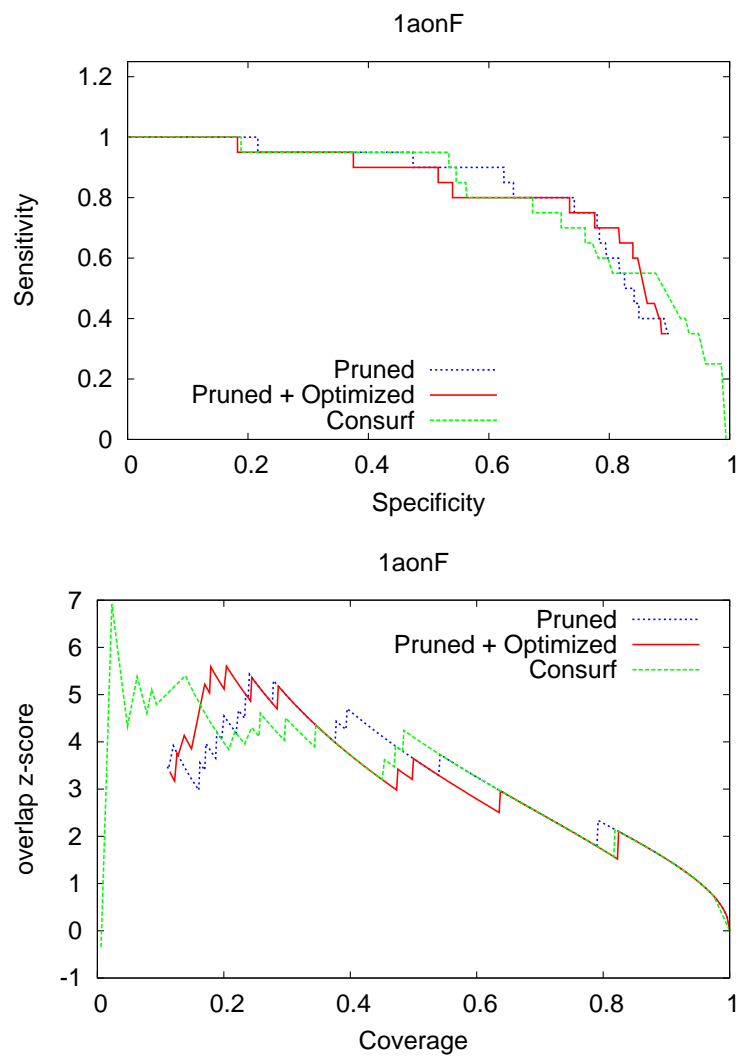
Figure 15: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a6q. b) The overlap z-scores as a function of coverage for the three methods is shown.
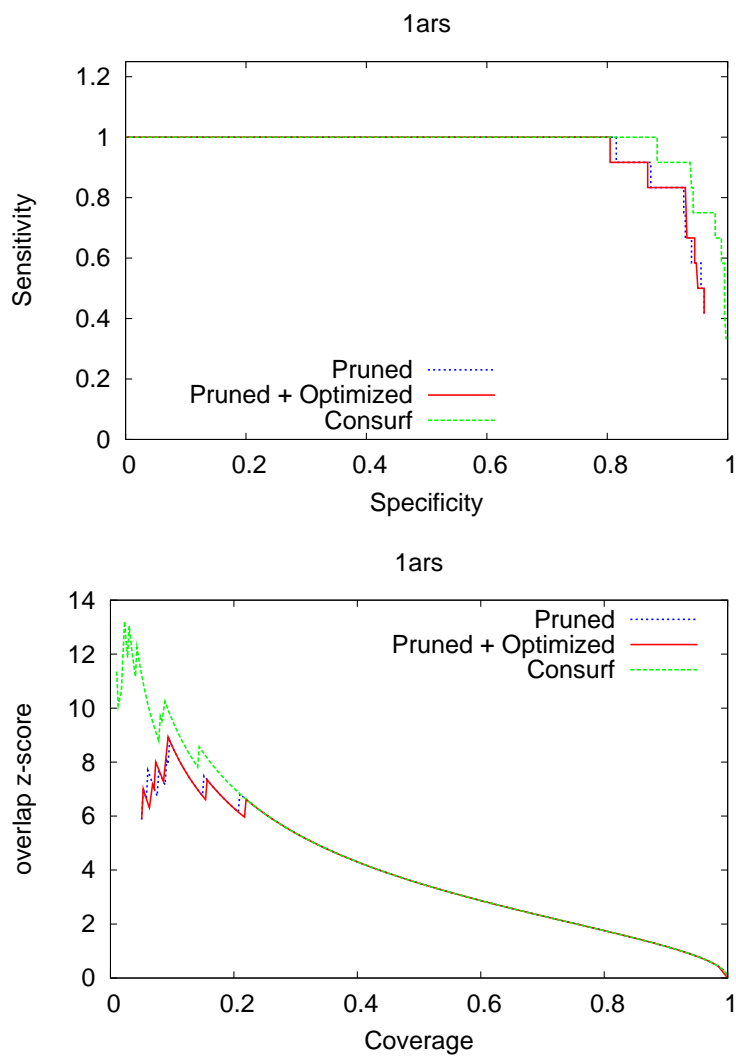
Figure 16: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1a80. b) The overlap z-scores as a function of coverage for the three methods is shown.
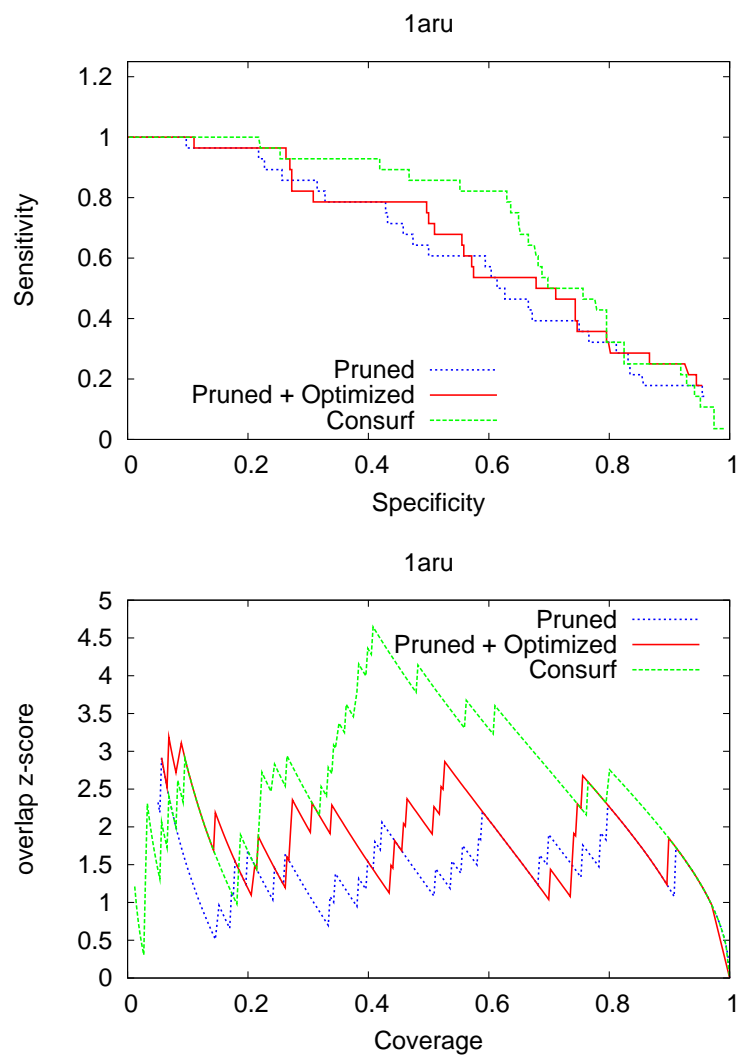
Figure 17: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aca. b) The overlap z-scores as a function of coverage for the three methods is shown.
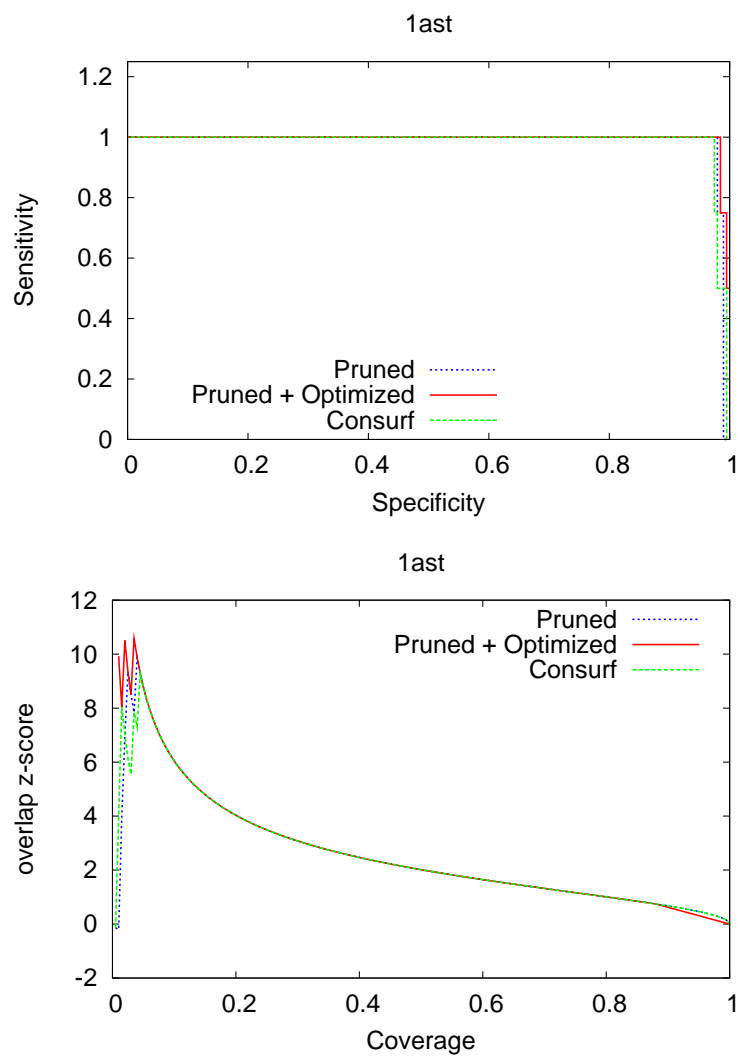
Figure 18: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ad3A. b) The overlap z-scores as a function of coverage for the three methods is shown.
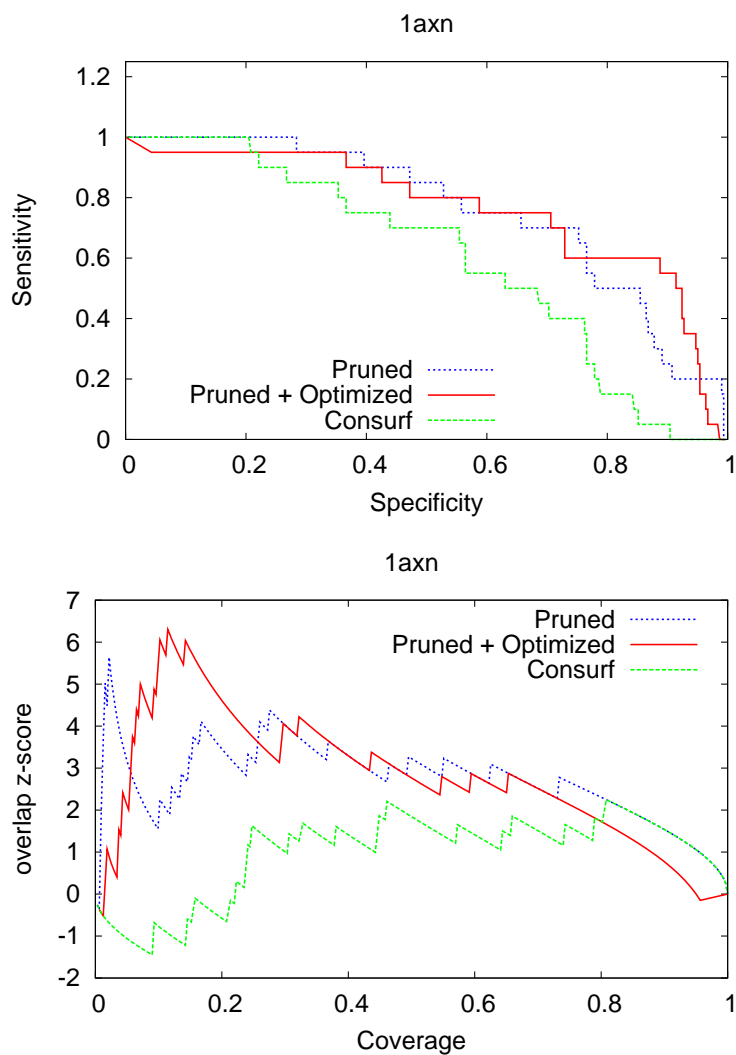
Figure 19: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ai2. b) The overlap z-scores as a function of coverage for the three methods is shown.
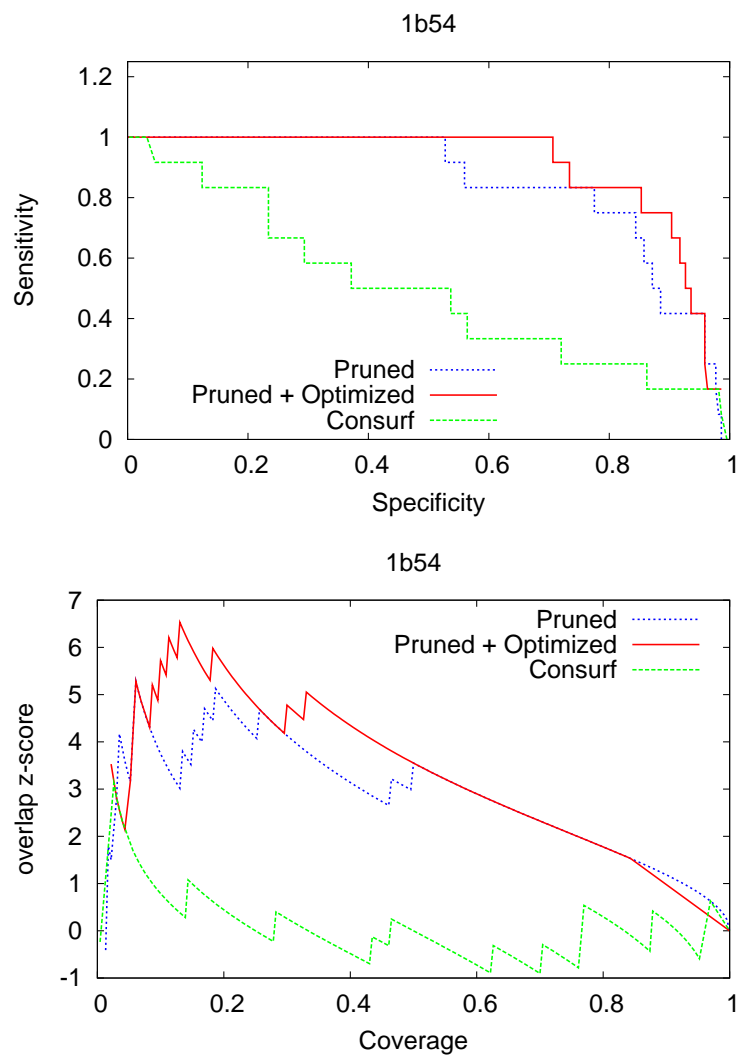
Figure 20: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aj2. b) The overlap z-scores as a function of coverage for the three methods is shown.
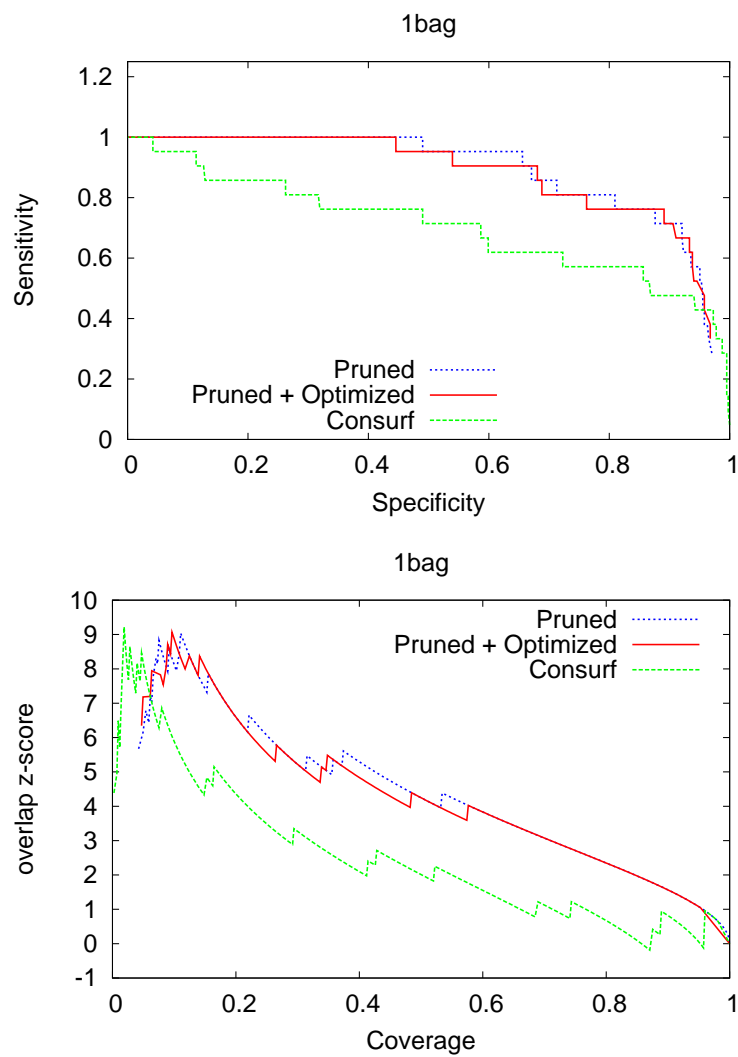
Figure 21: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aj8A. b) The overlap z-scores as a function of coverage for the three methods is shown.
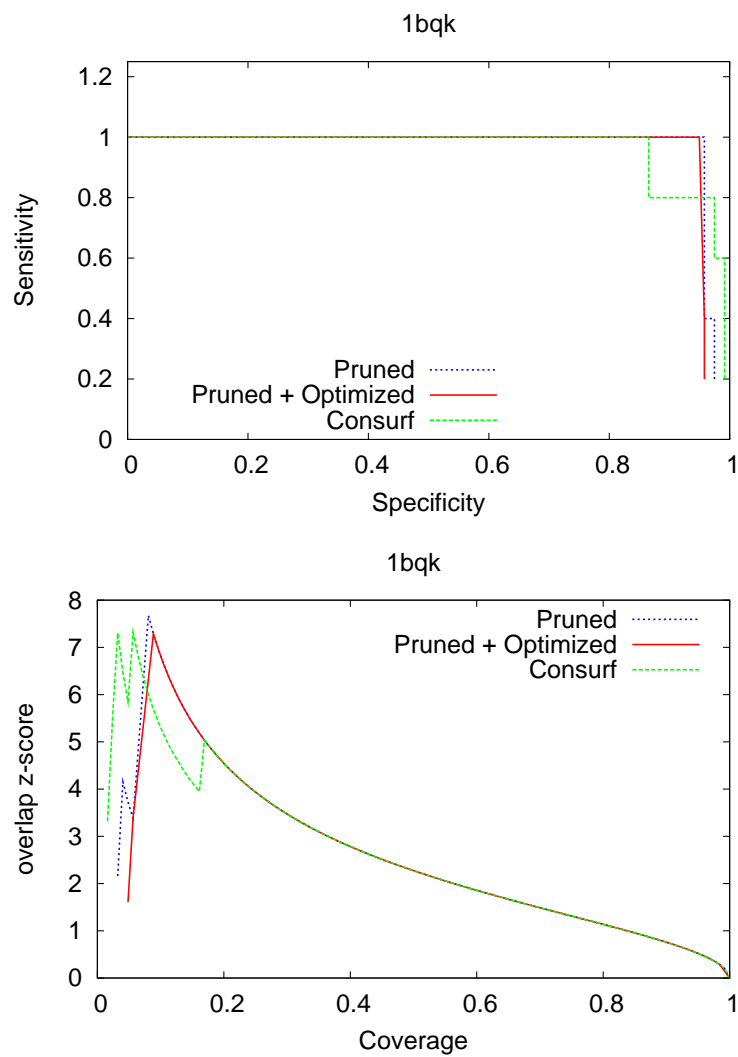
Figure 22: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aky. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 23: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1am1. b) The overlap z-scores as a function of coverage for the three methods is shown.
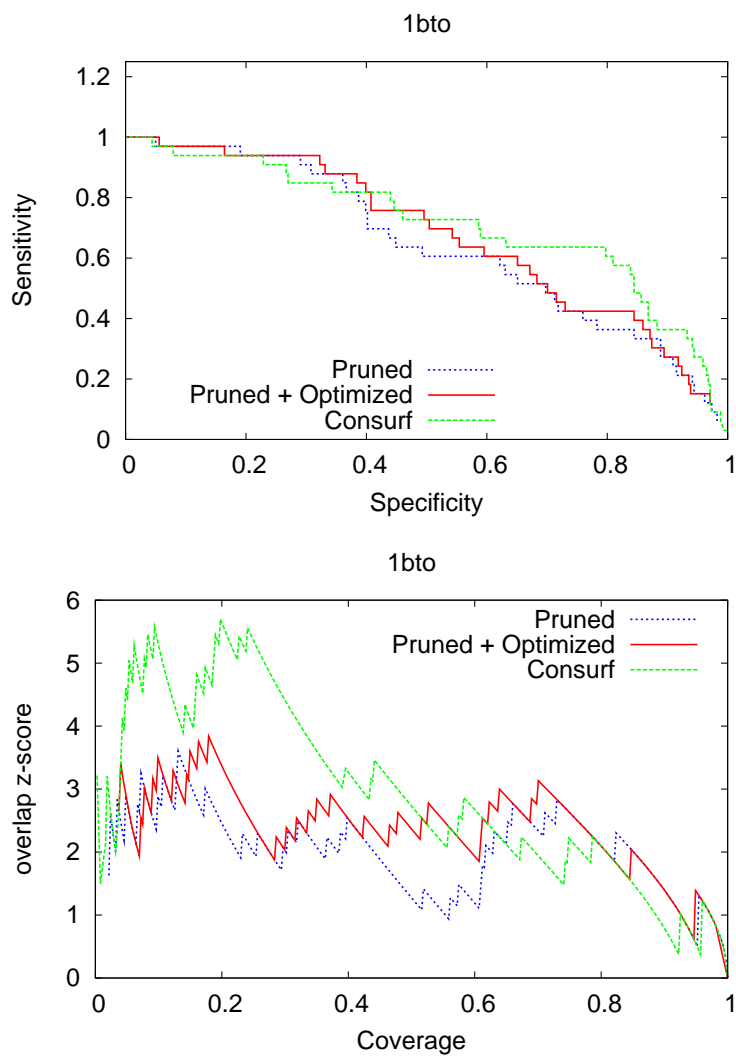
Figure 24: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1amk. b) The overlap z-scores as a function of coverage for the three methods is shown.
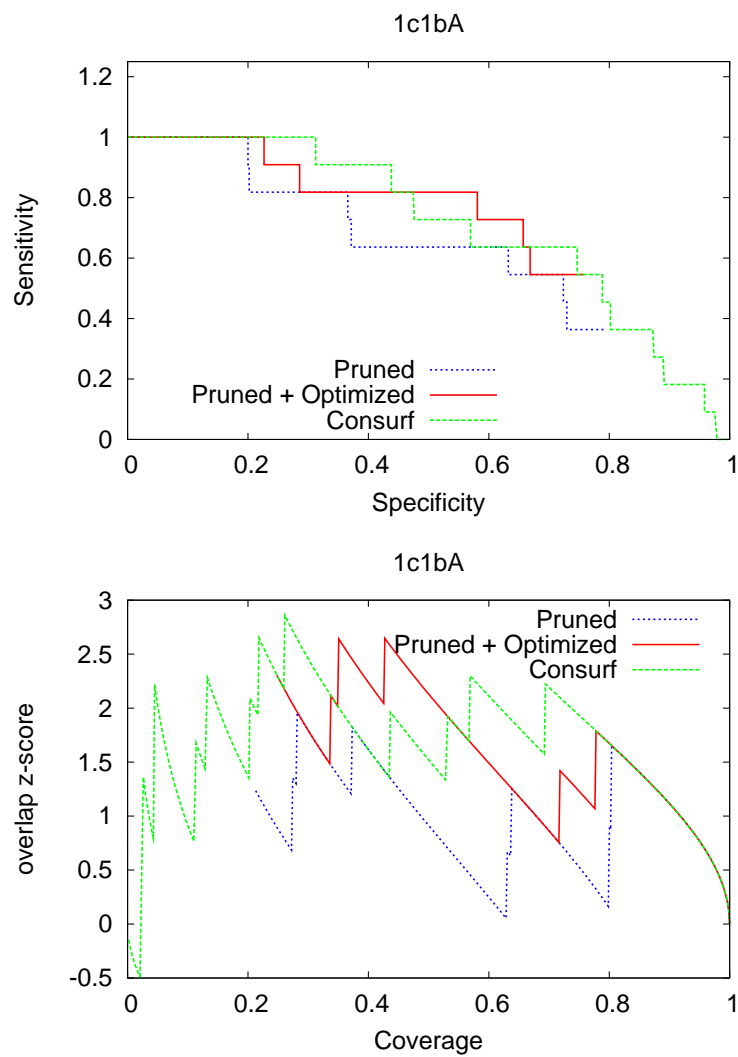
Figure 25: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aonF. b) The overlap z-scores as a function of coverage for the three methods is shown.
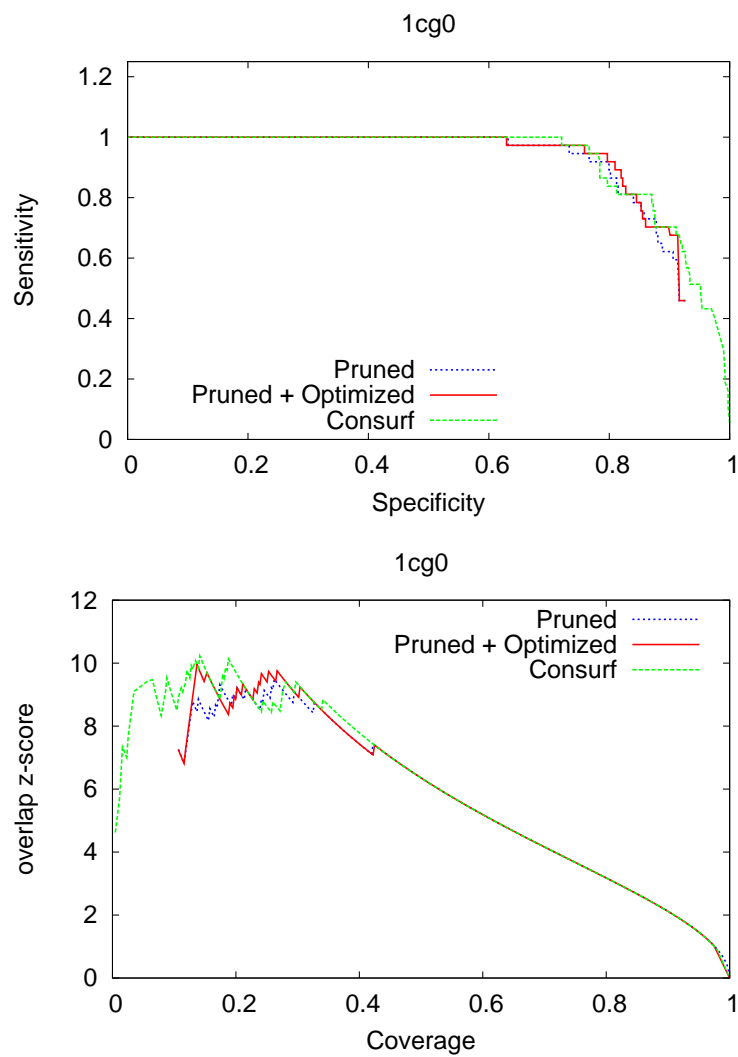
Figure 26: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ars. b) The overlap z-scores as a function of coverage for the three methods is shown.
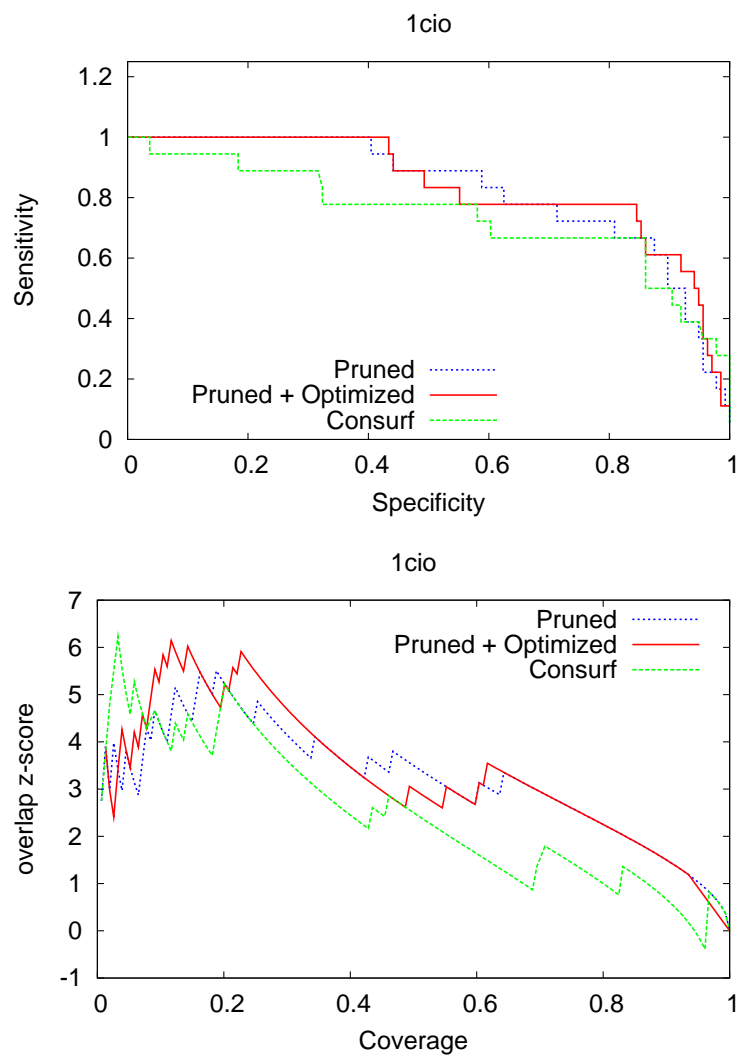
Figure 27: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aru. b) The overlap z-scores as a function of coverage for the three methods is shown.
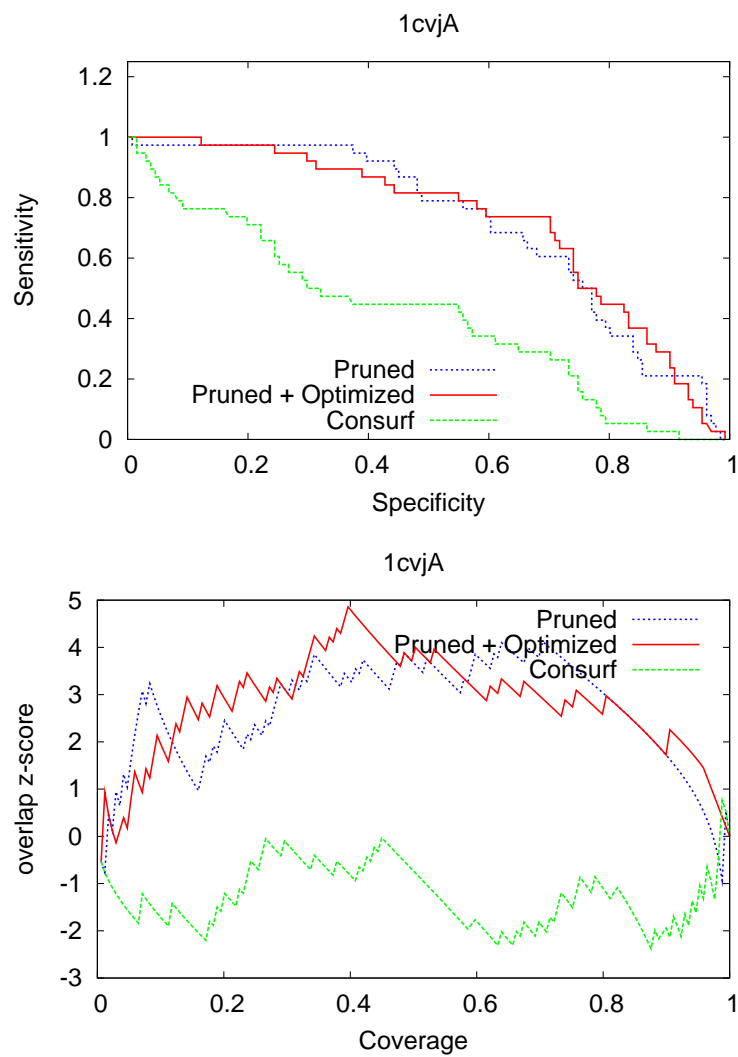
Figure 28: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ast. b) The overlap z-scores as a function of coverage for the three methods is shown.
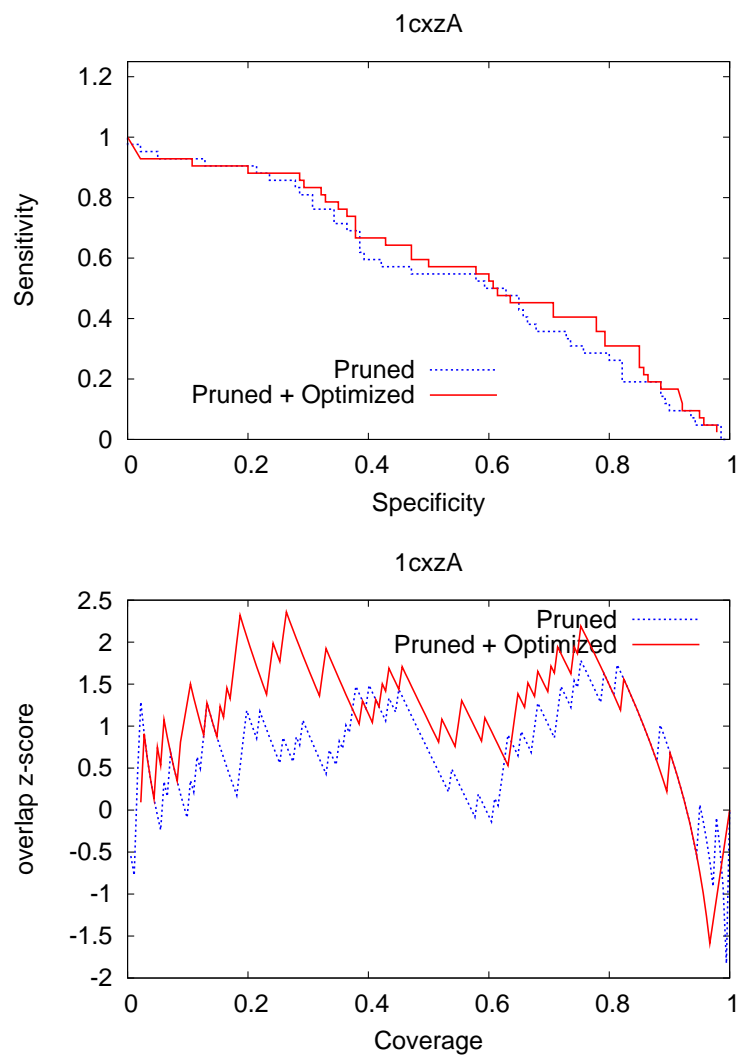
Figure 29: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1axn. b) The overlap z-scores as a function of coverage for the three methods is shown.
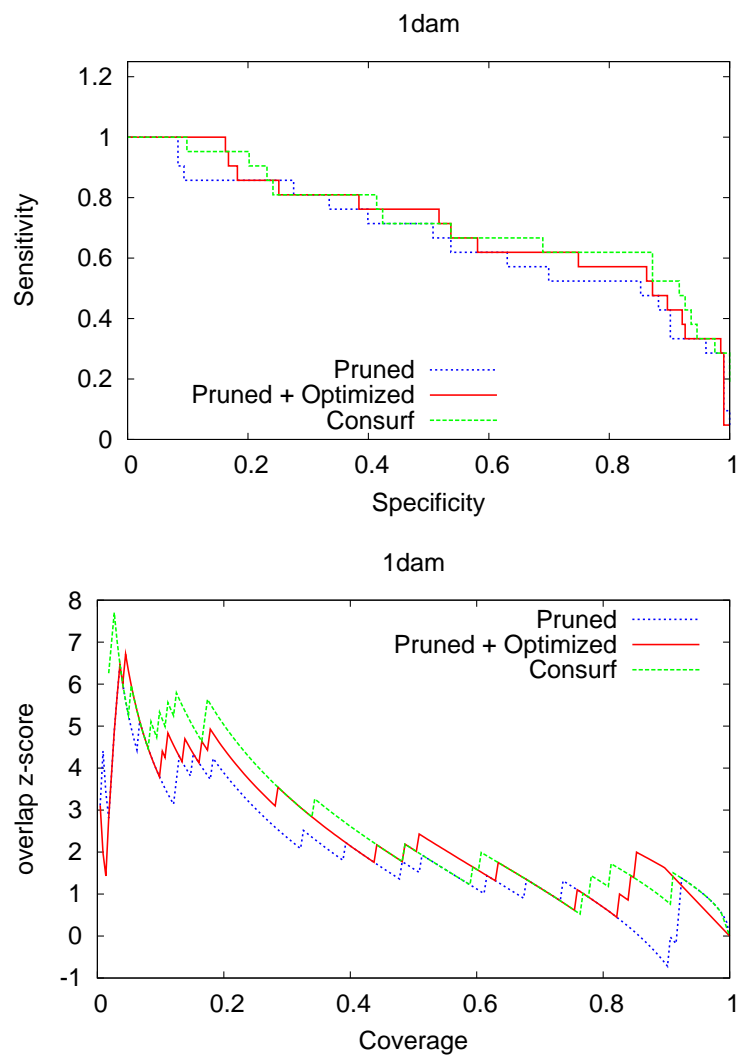
Figure 30: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1b54. b) The overlap z-scores as a function of coverage for the three methods is shown.
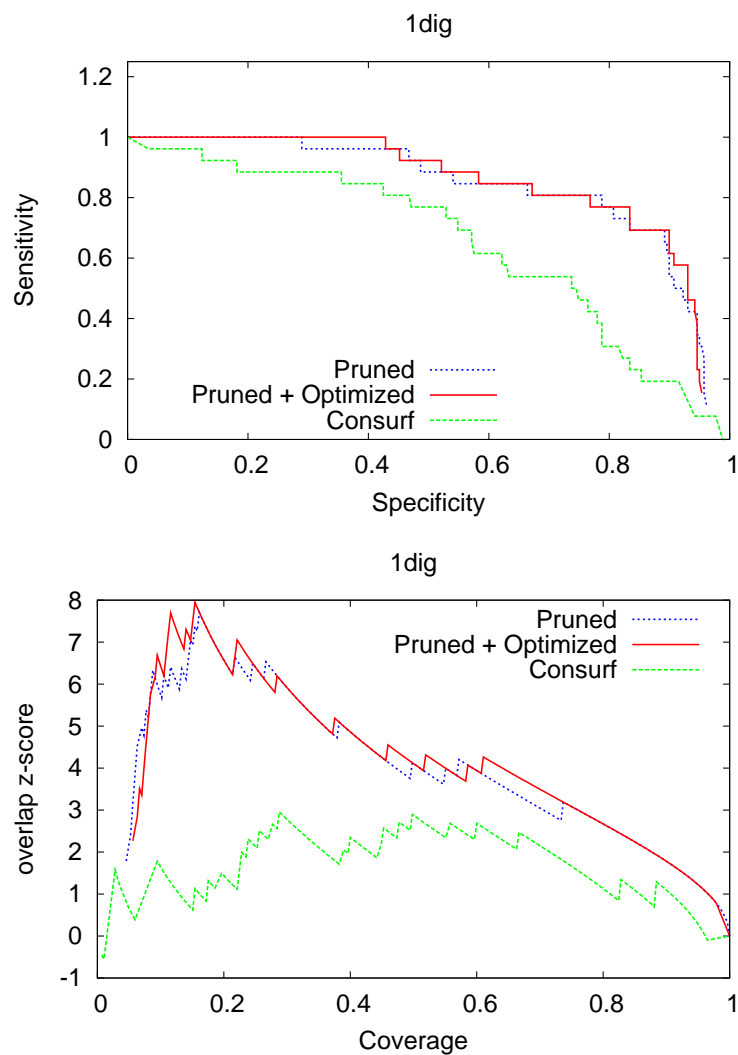
Figure 31: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bag. b) The overlap z-scores as a function of coverage for the three methods is shown.
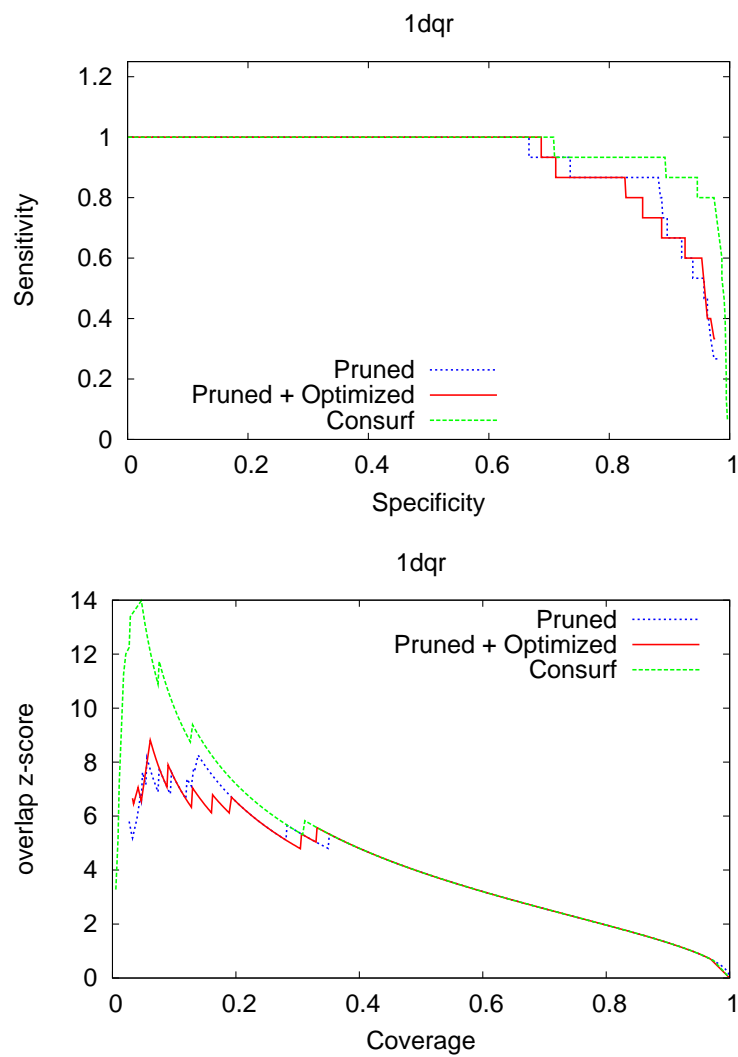
Figure 32: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bqk. b) The overlap z-scores as a function of coverage for the three methods is shown.
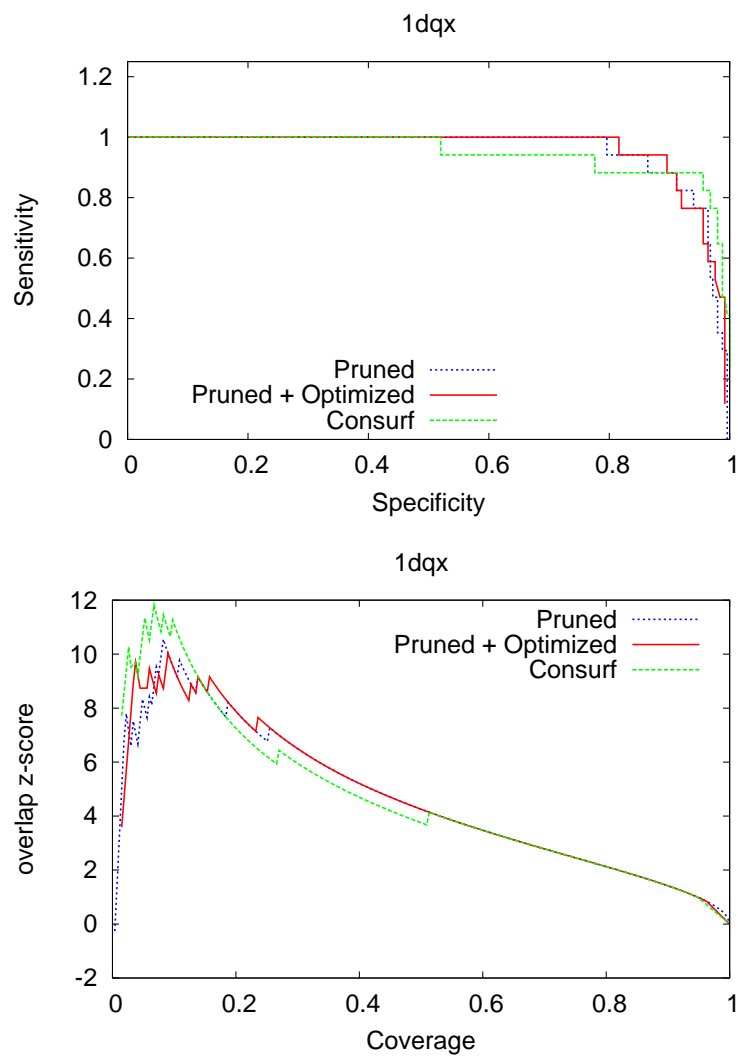
Figure 33: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bto. b) The overlap z-scores as a function of coverage for the three methods is shown.
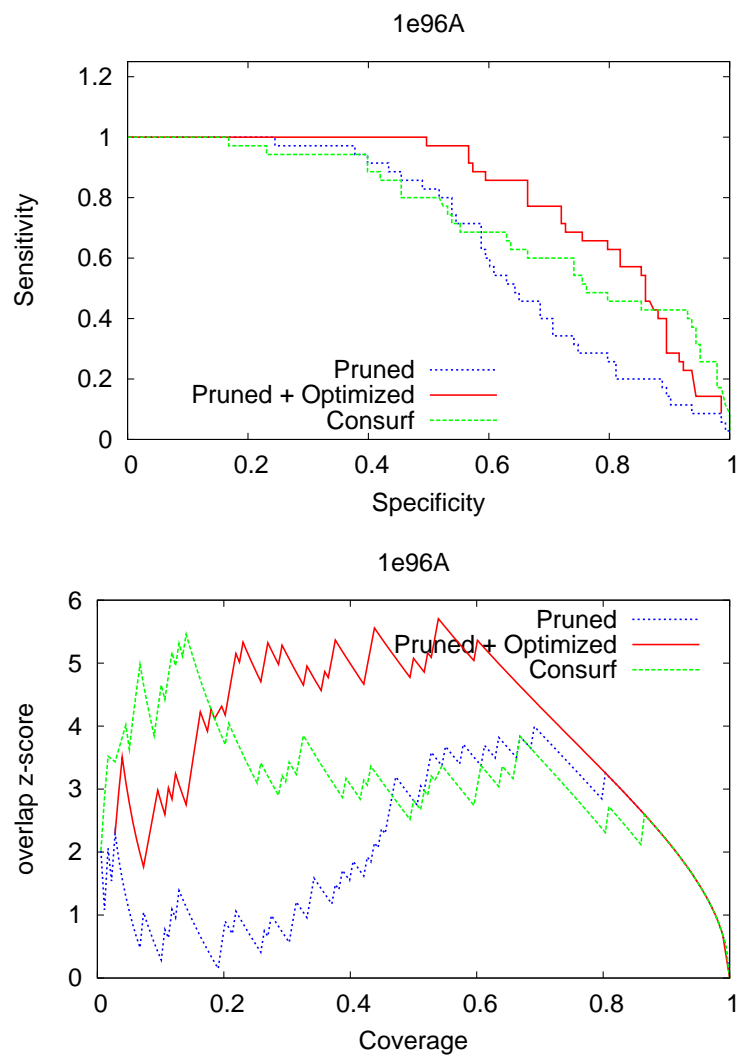
Figure 34: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1c1bA. b) The overlap z-scores as a function of coverage for the three methods is shown.
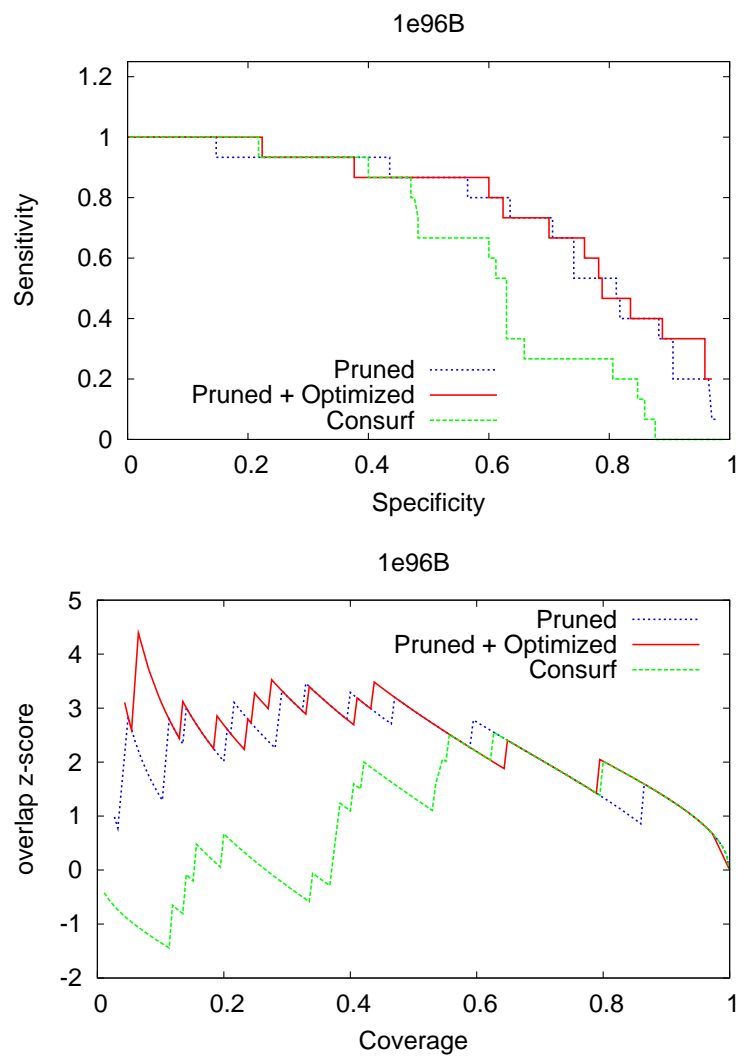
Figure 35: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1cg0. b) The overlap z-scores as a function of coverage for the three methods is shown.
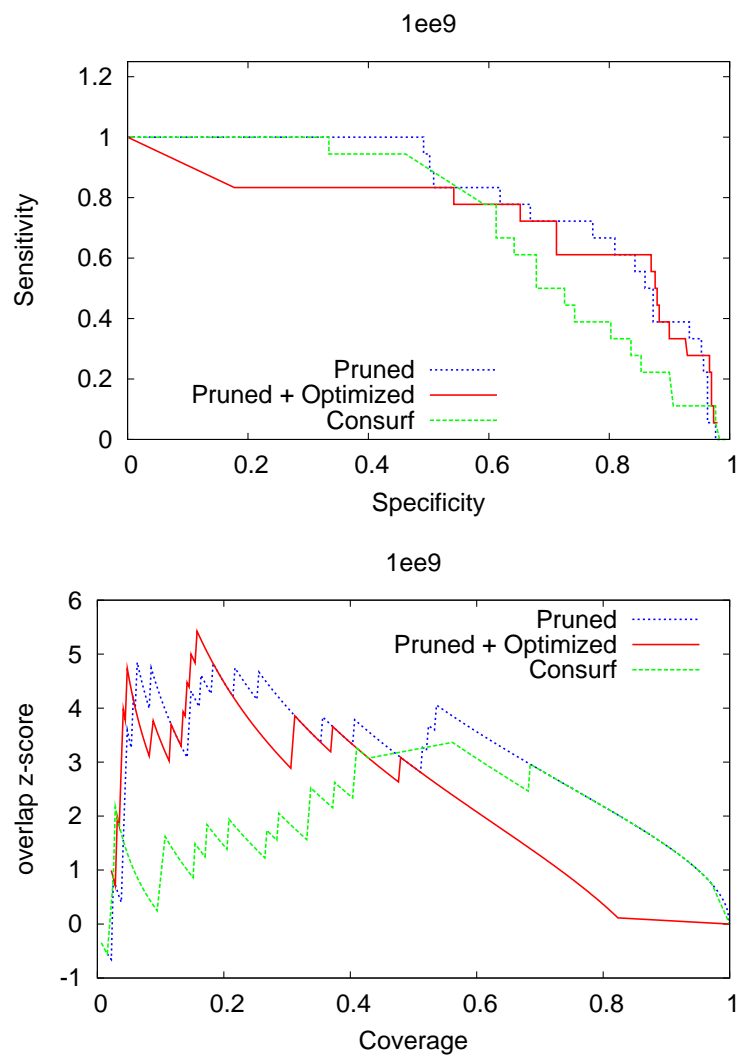
Figure 36: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1cio. b) The overlap z-scores as a function of coverage for the three methods is shown.
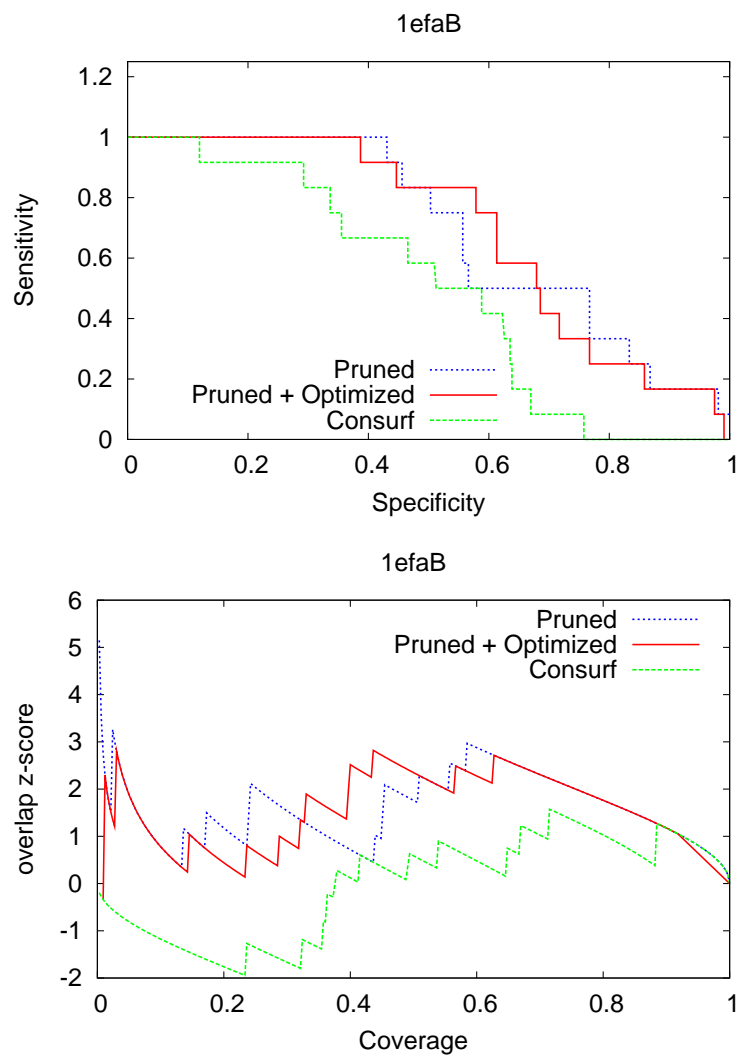
39

Figure 37: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1cvjA. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 38: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1cxzA. b) The overlap z-scores as a function of coverage for the three methods is shown.
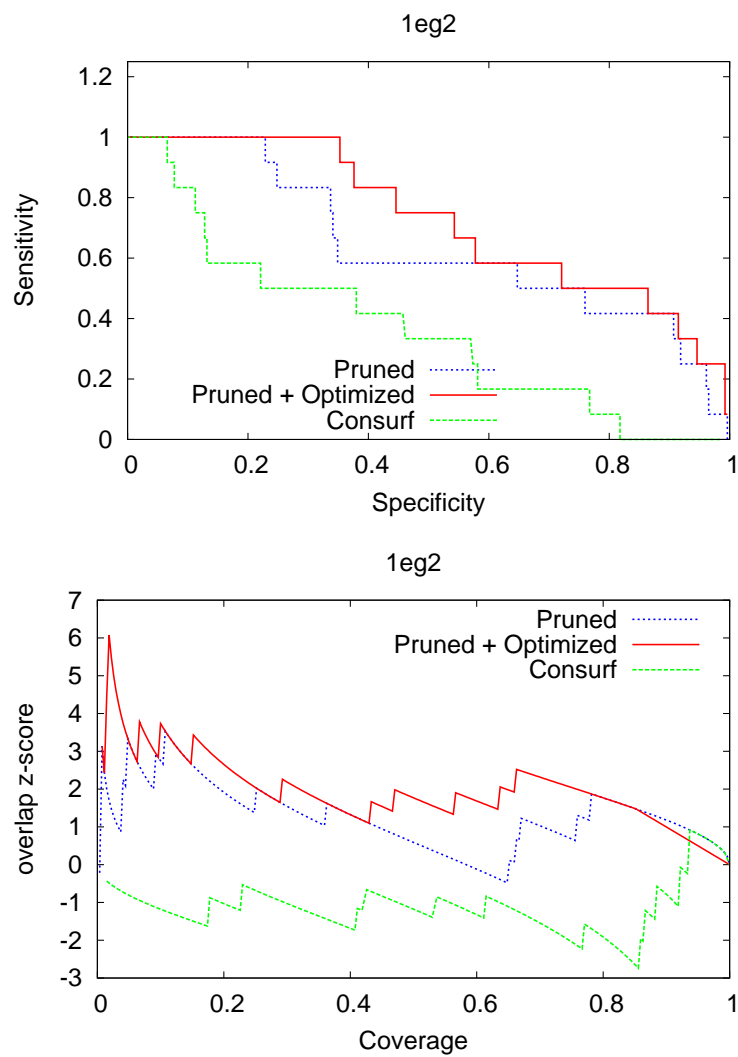
Figure 39: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1dam. b) The overlap z-scores as a function of coverage for the three methods is shown.
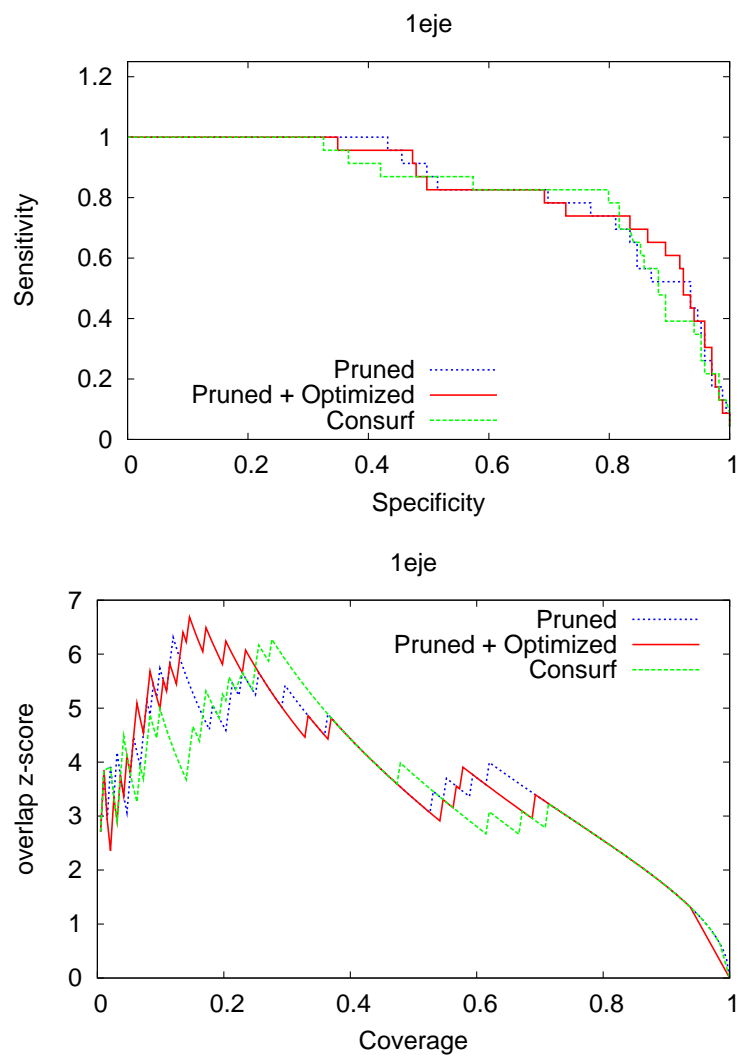
Figure 40: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1dig. b) The overlap z-scores as a function of coverage for the three methods is shown.
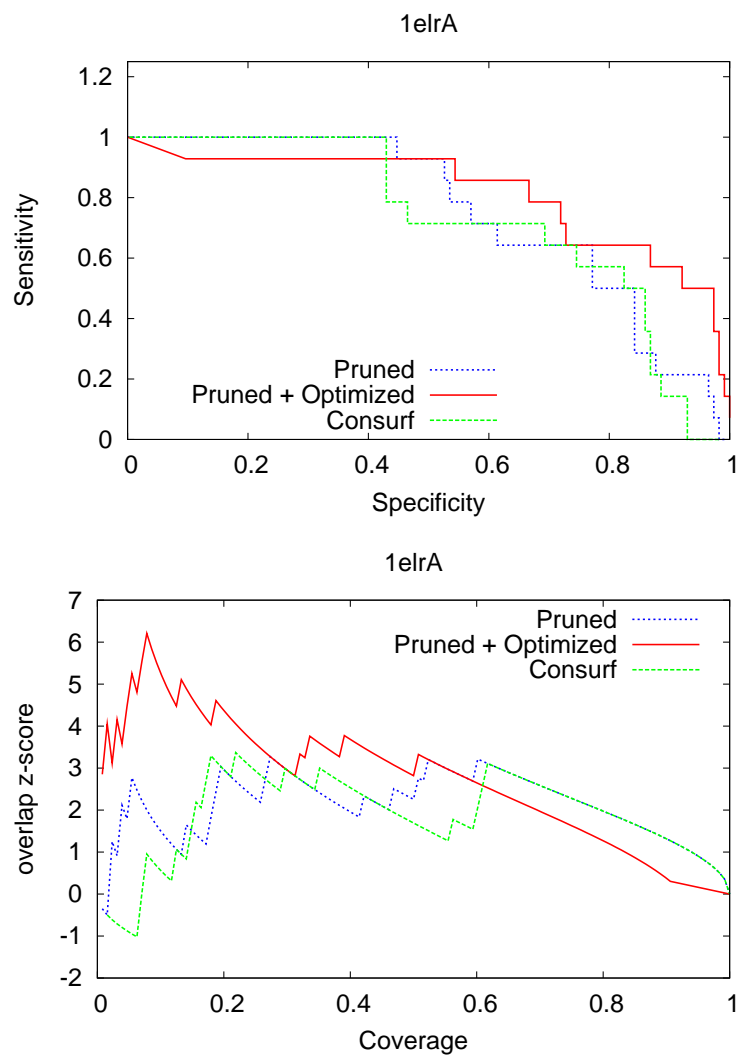
Figure 41: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1dqr. b) The overlap z-scores as a function of coverage for the three methods is shown.
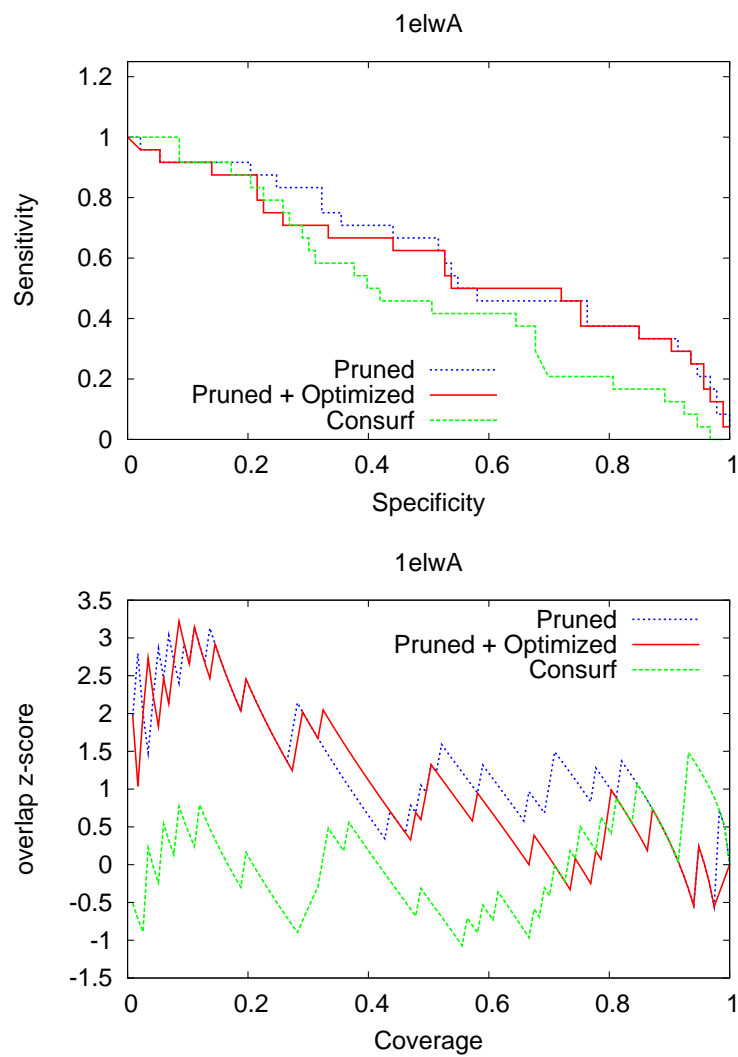
Figure 42: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1dqx. b) The overlap z-scores as a function of coverage for the three methods is shown.
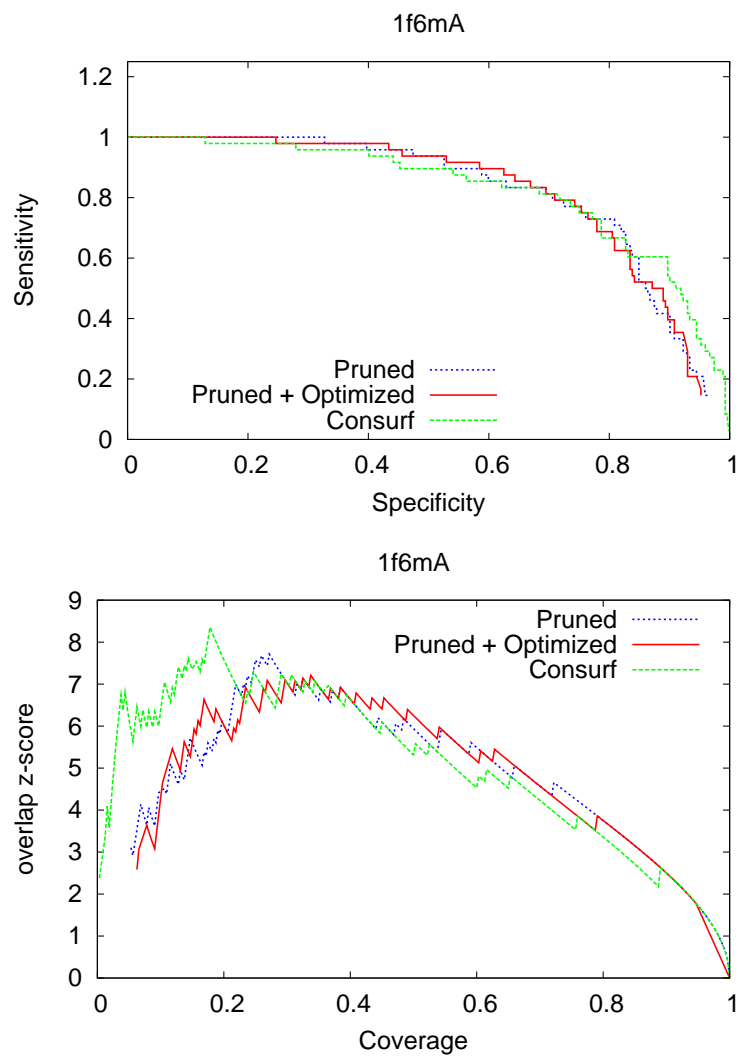
Figure 43: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1e96A. b) The overlap z-scores as a function of coverage for the three methods is shown.
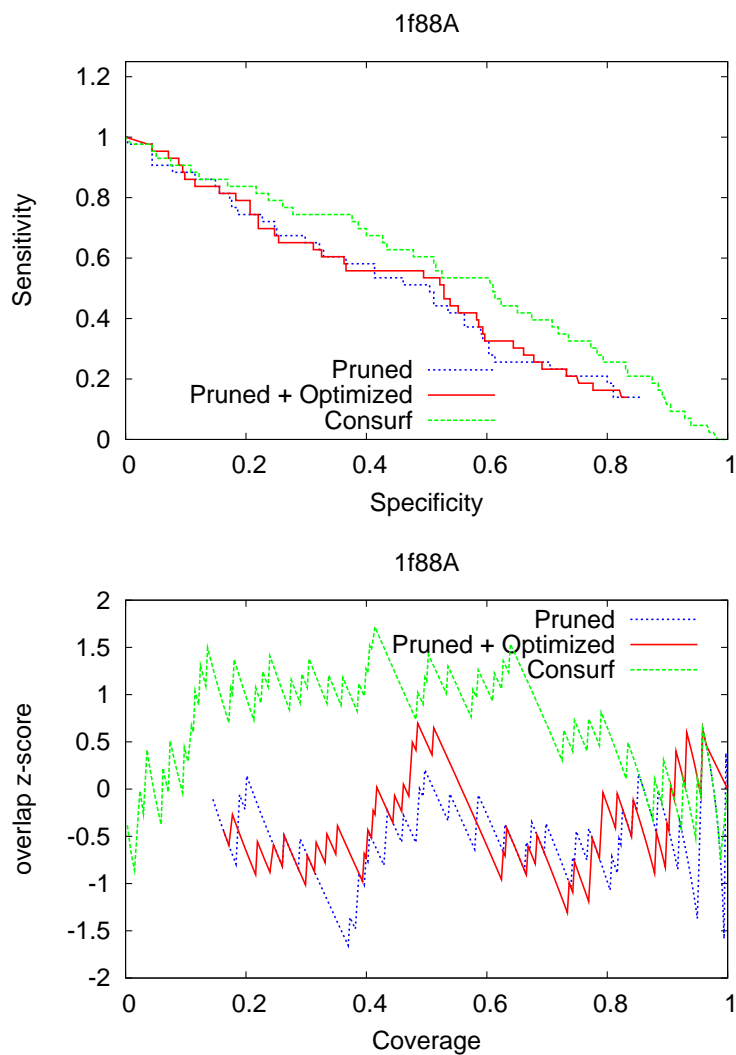
Figure 44: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1e96B. b) The overlap z-scores as a function of coverage for the three methods is shown.
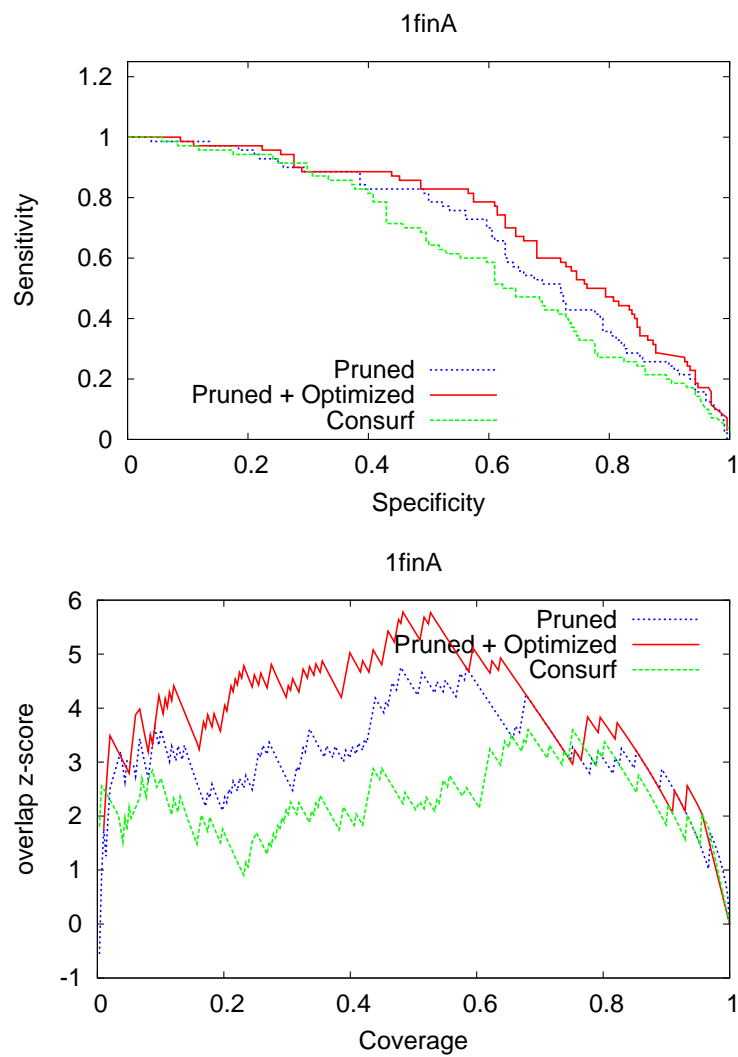
Figure 45: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ee9. b) The overlap z-scores as a function of coverage for the three methods is shown.
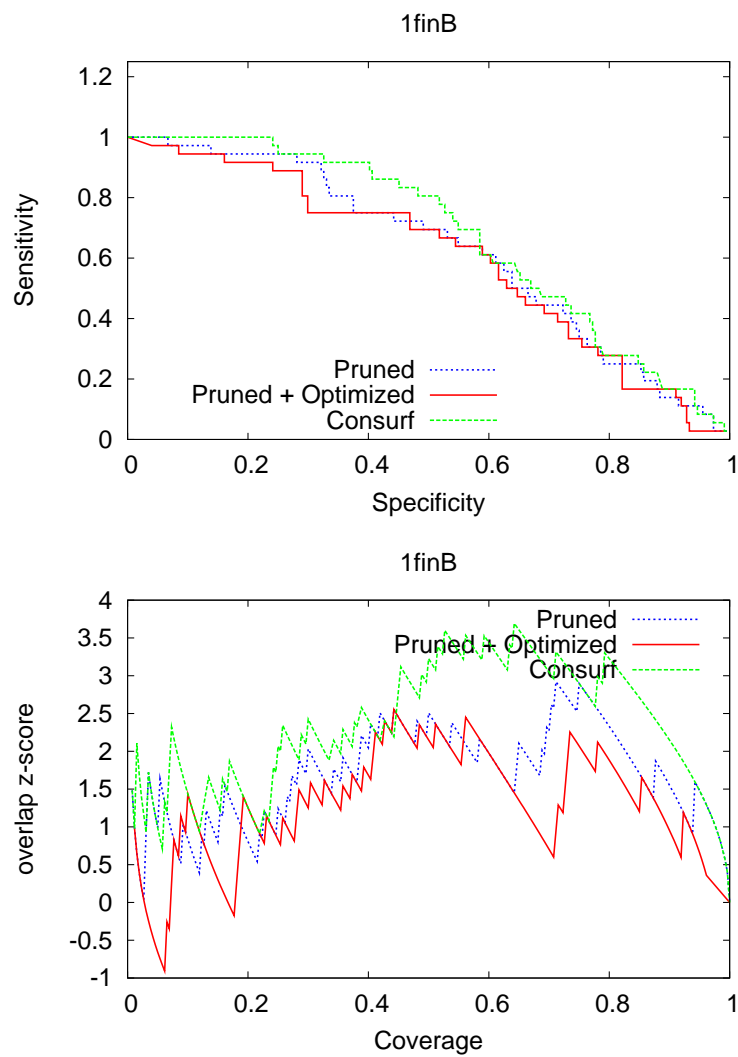
Figure 46: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1efaB. b) The overlap z-scores as a function of coverage for the three methods is shown.
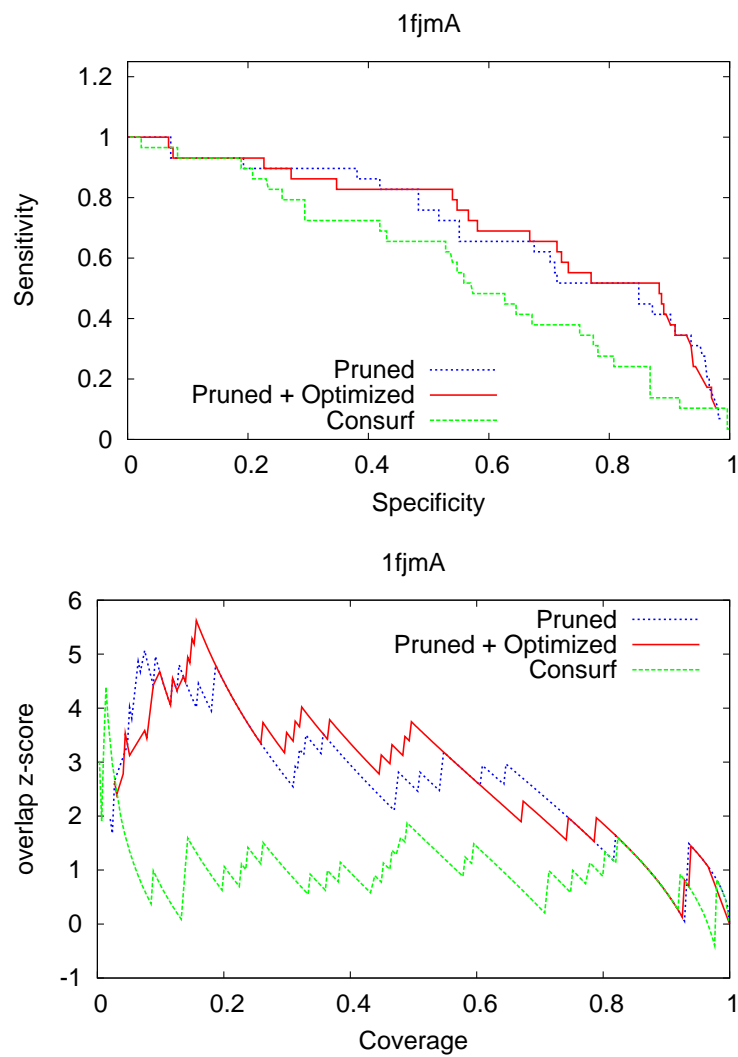
Figure 47: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1eg2. b) The overlap z-scores as a function of coverage for the three methods is shown.
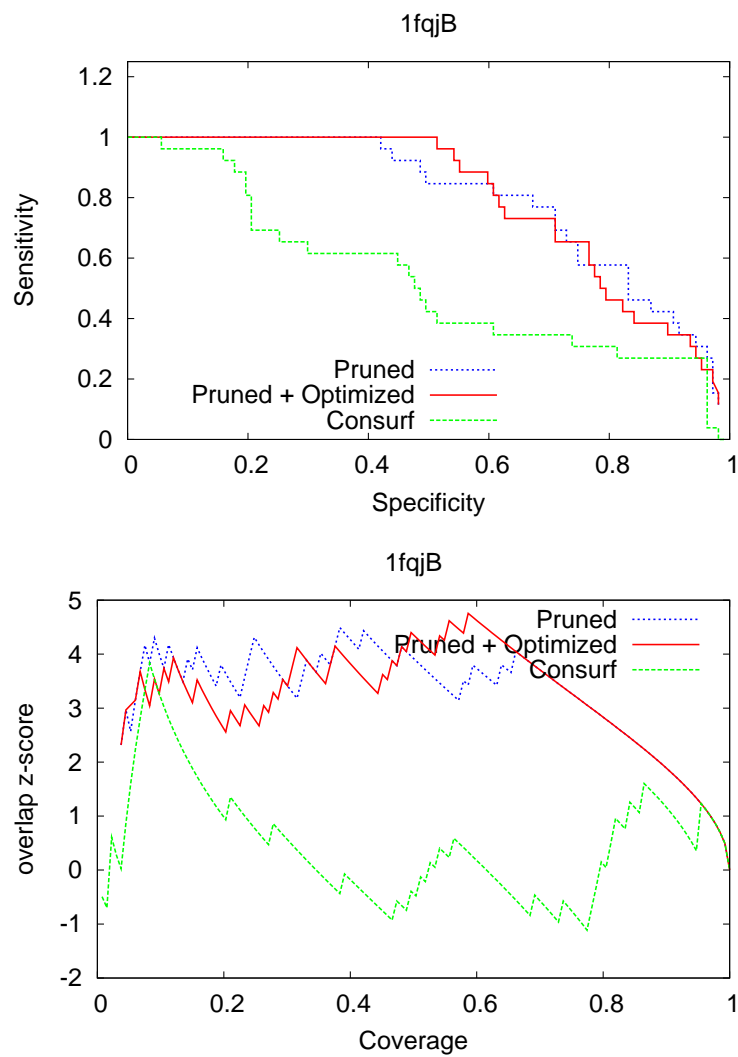
Figure 48: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1eje. b) The overlap z-scores as a function of coverage for the three methods is shown.
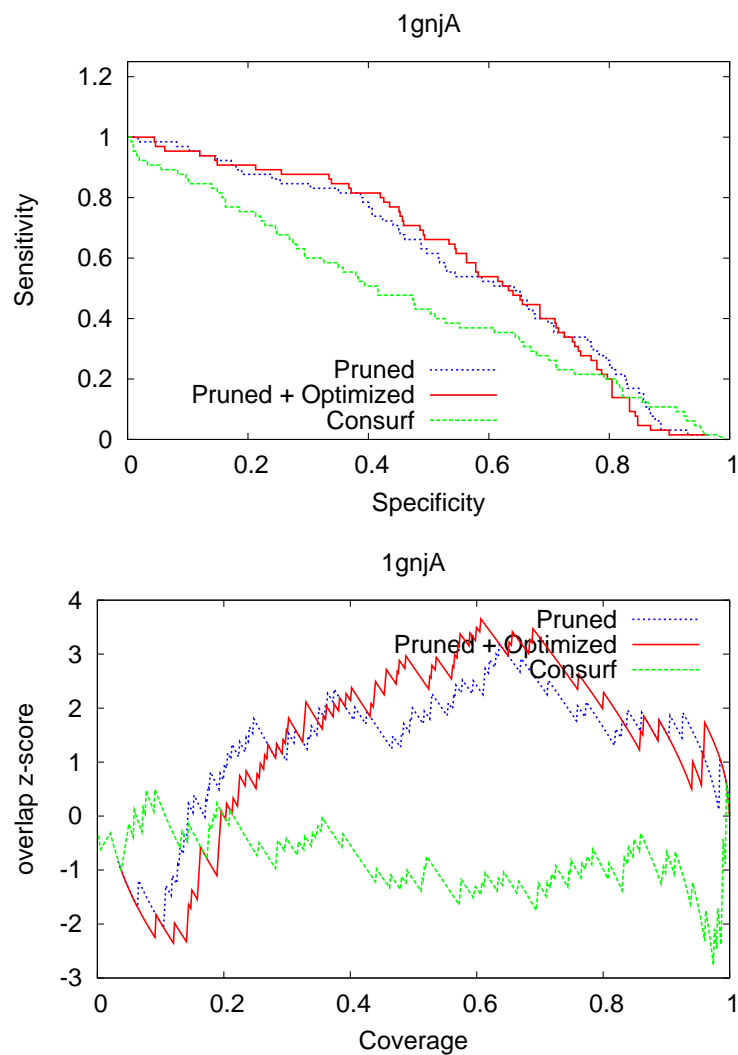
Figure 49: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1elrA. b) The overlap z-scores as a function of coverage for the three methods is shown.
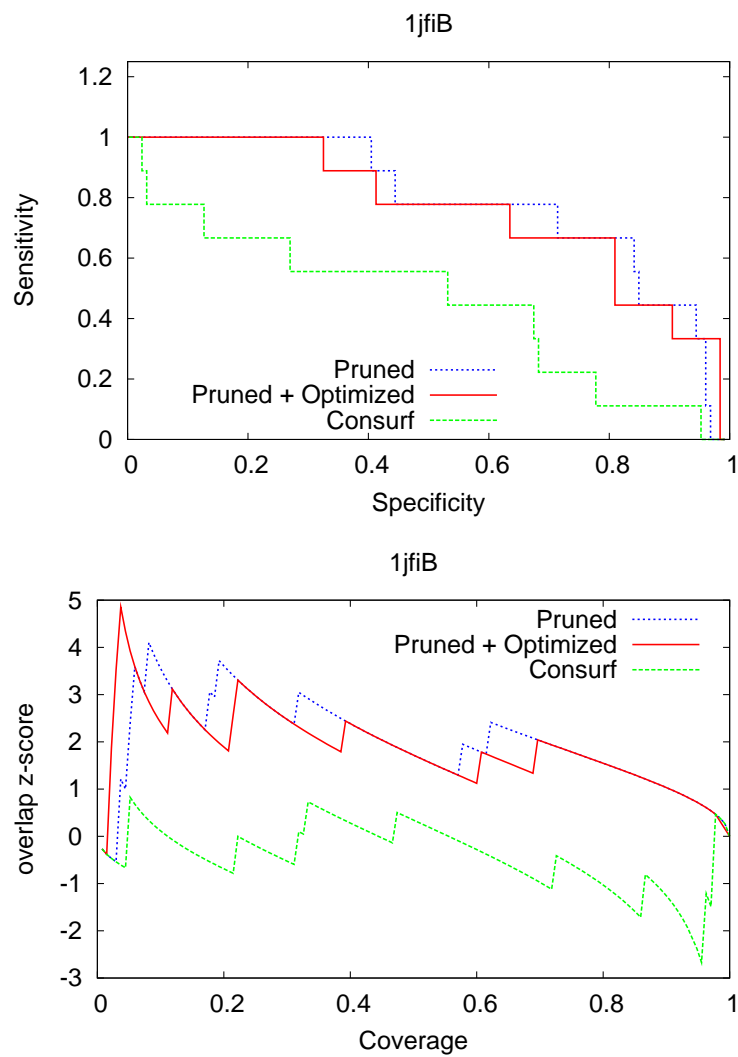
Figure 50: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1elwA. b) The overlap z-scores as a function of coverage for the three methods is shown.
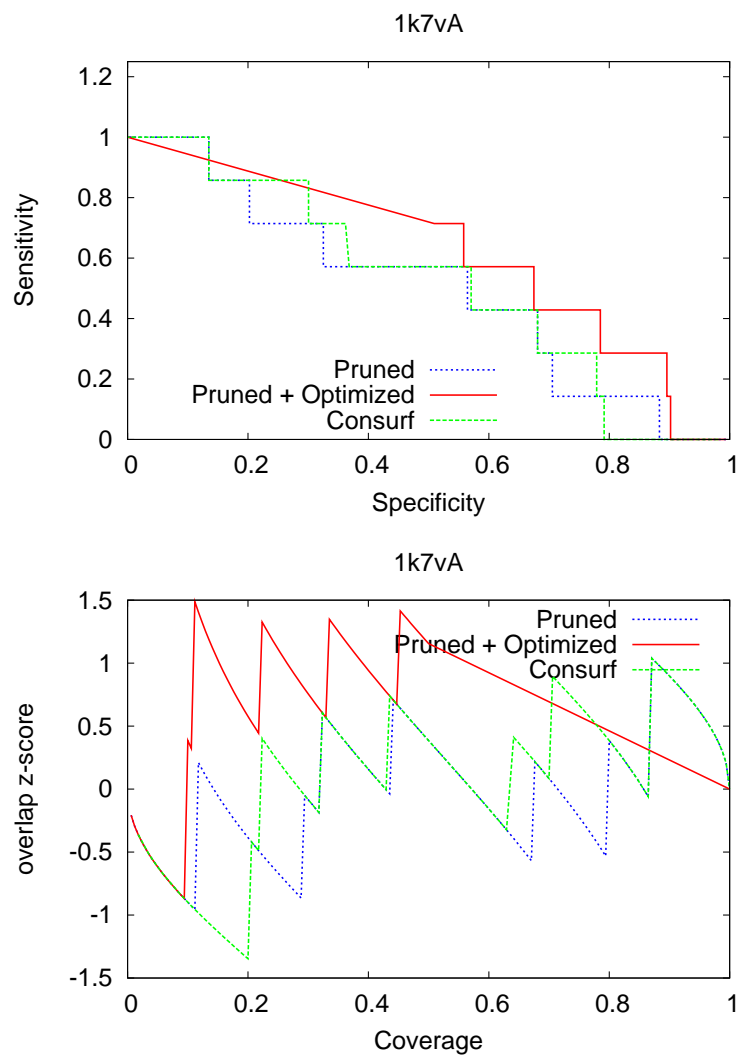
Figure 51: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1f6mA. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 52: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1f88A. b) The overlap z-scores as a function of coverage for the three methods is shown.
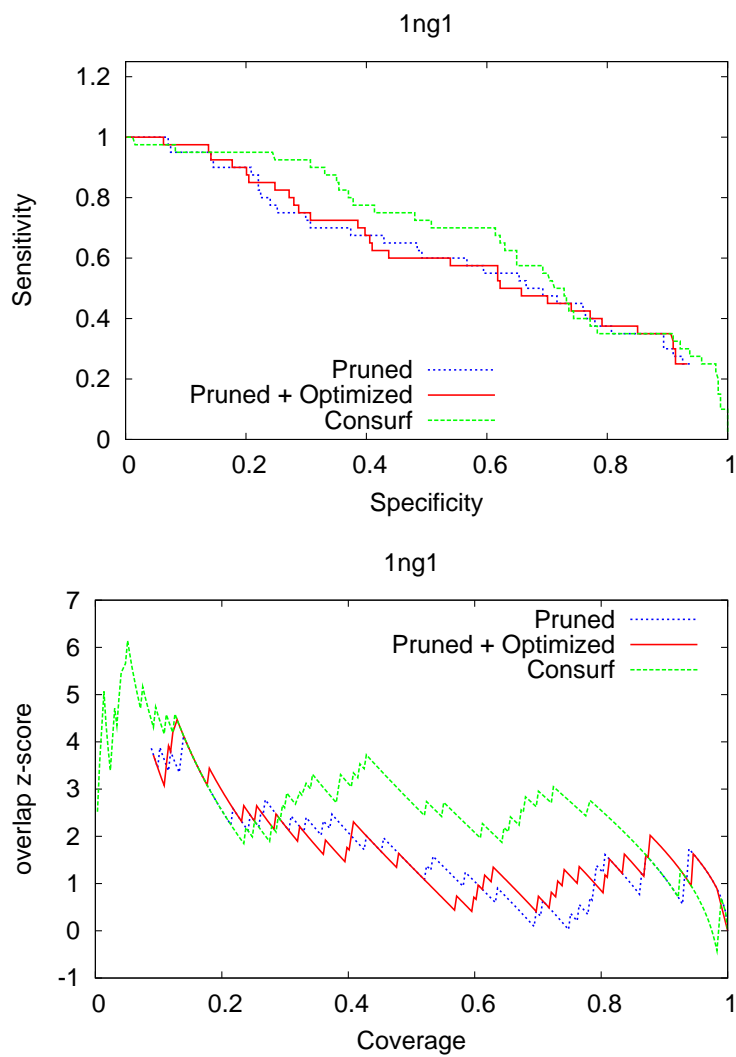
Figure 53:   a) The sensitivity and specifity is shown for the pruned, pruned
+ optimized and the comparison method Consurf for PDBID 1finA. b) The
overlap z-scores as a function of coverage for the three methods is shown.
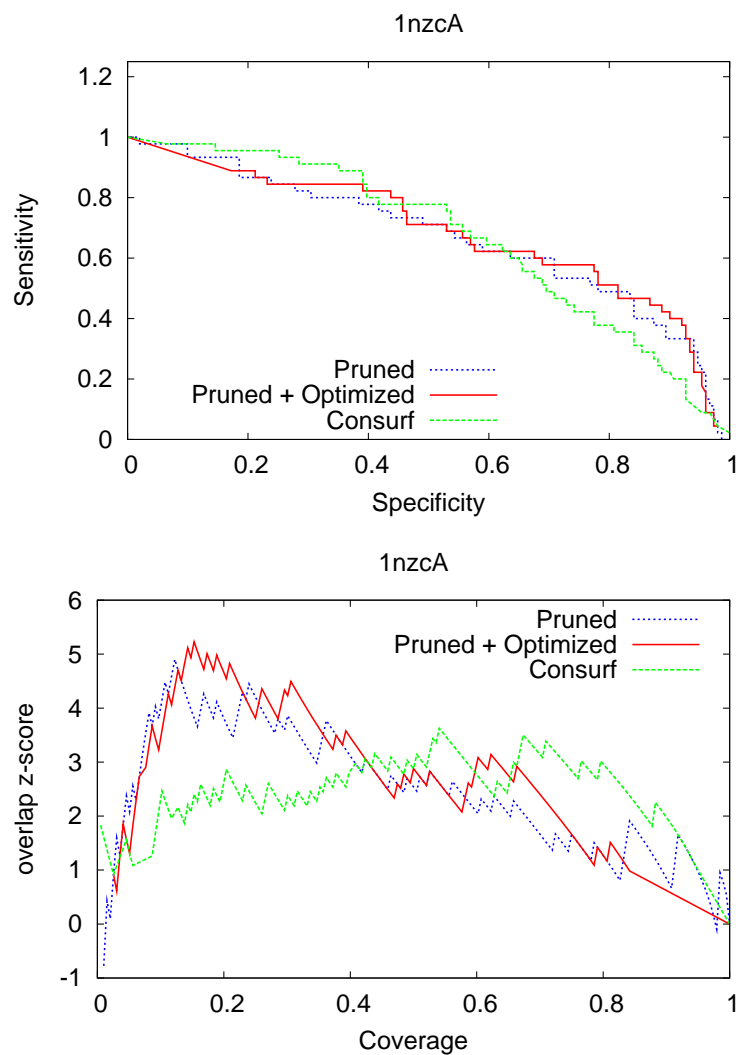
Figure 54: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1finB. b) The overlap z-scores as a function of coverage for the three methods is shown.
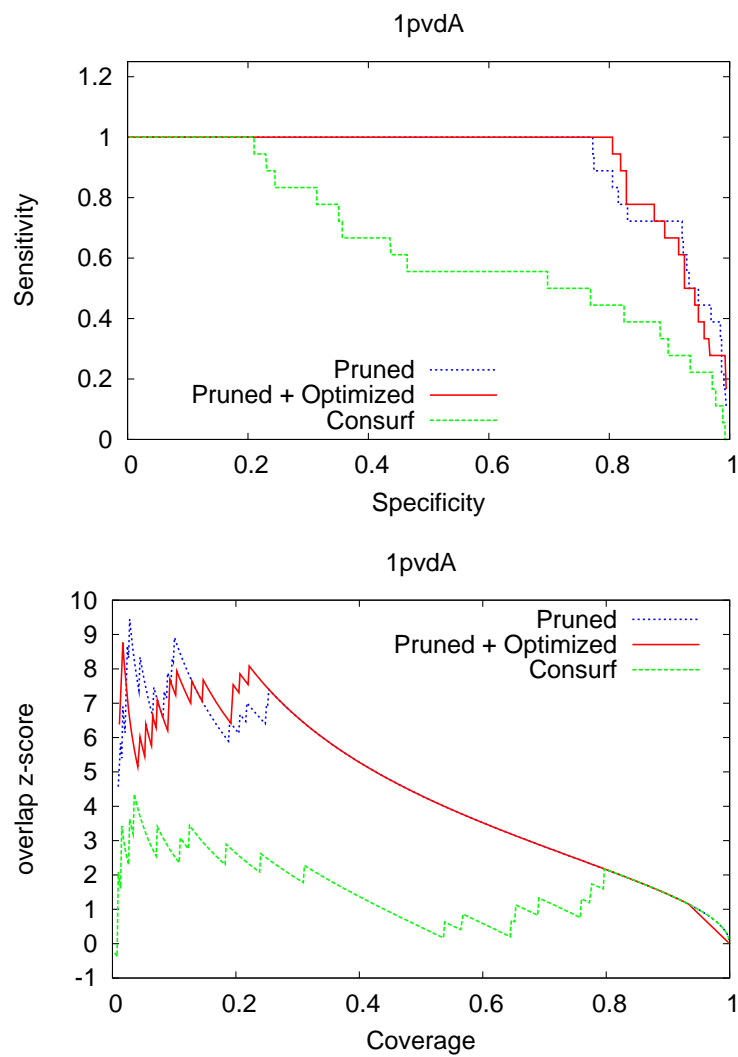
Figure 55: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fjmA. b) The overlap z-scores as a function of coverage for the three methods is shown.
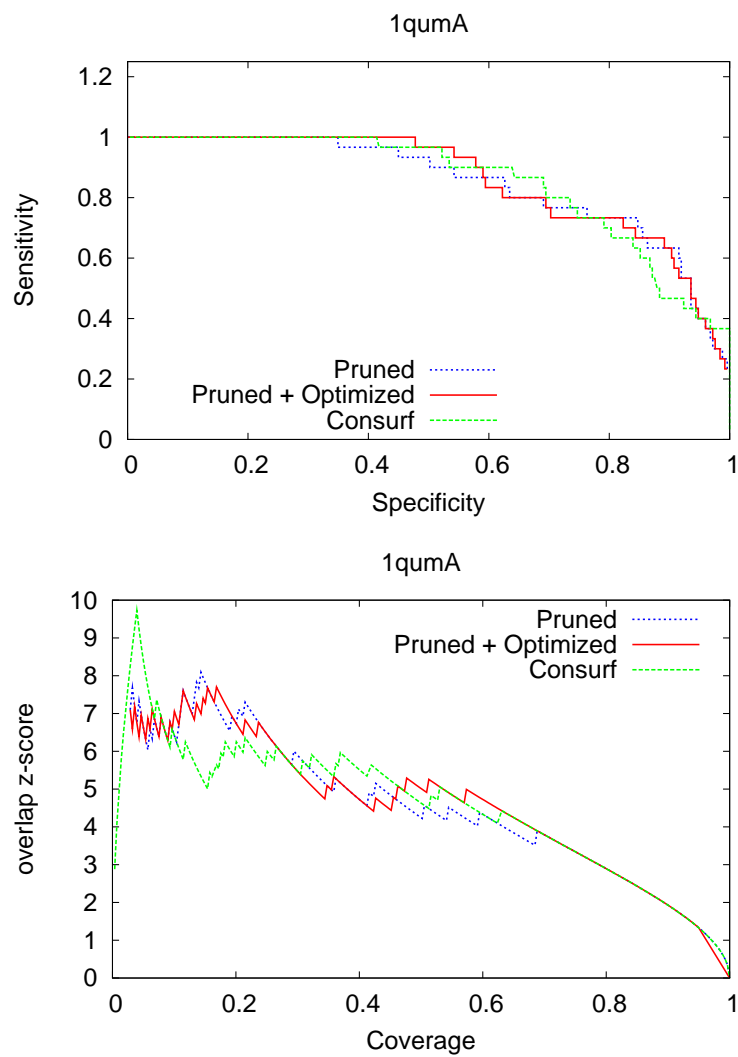
Figure 56: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fqjB. b) The overlap z-scores as a function of coverage for the three methods is shown.
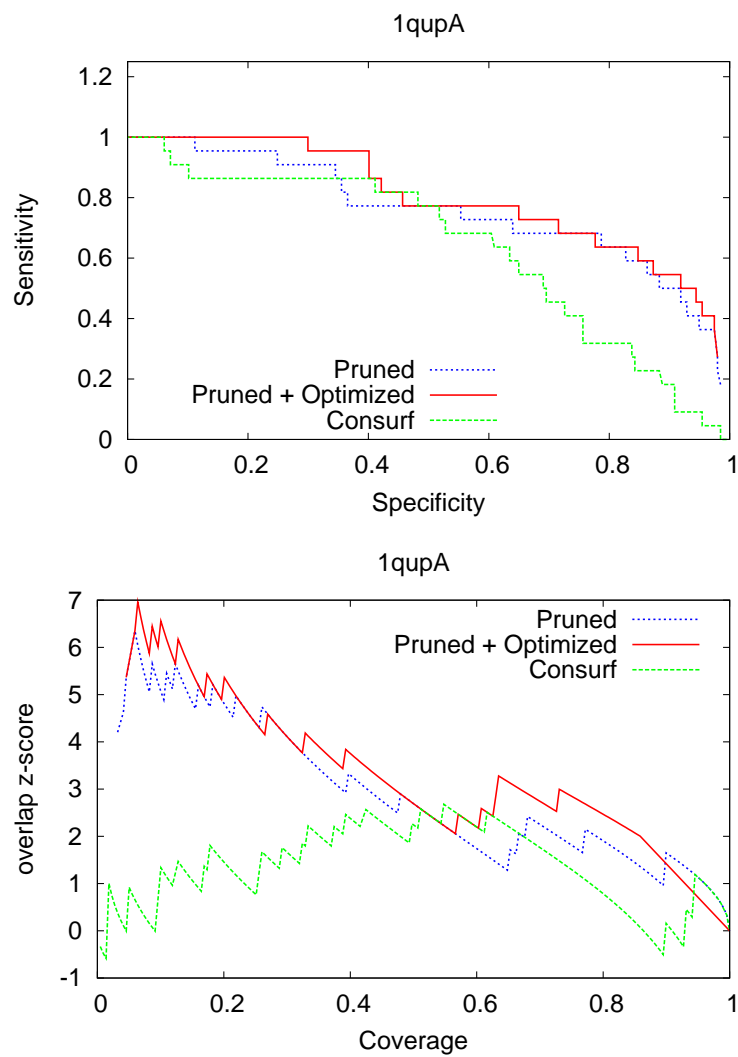
Figure 57: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1gnjA. b) The overlap z-scores as a function of coverage for the three methods is shown.
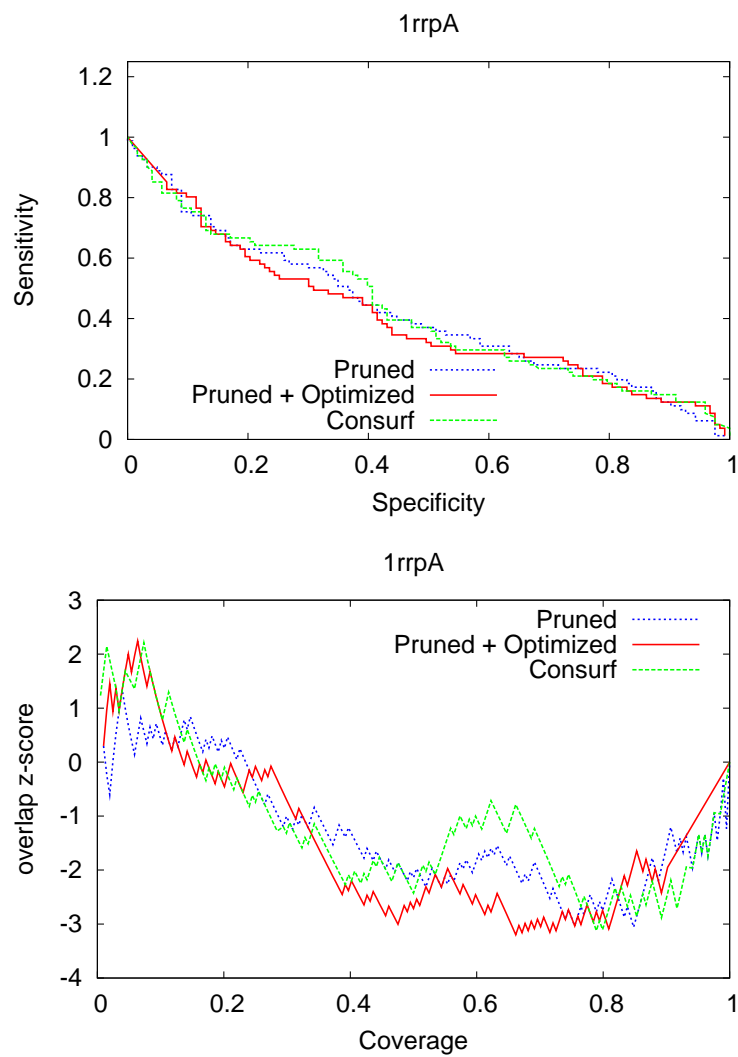
Figure 58: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1jfiB. b) The overlap z-scores as a function of coverage for the three methods is shown.
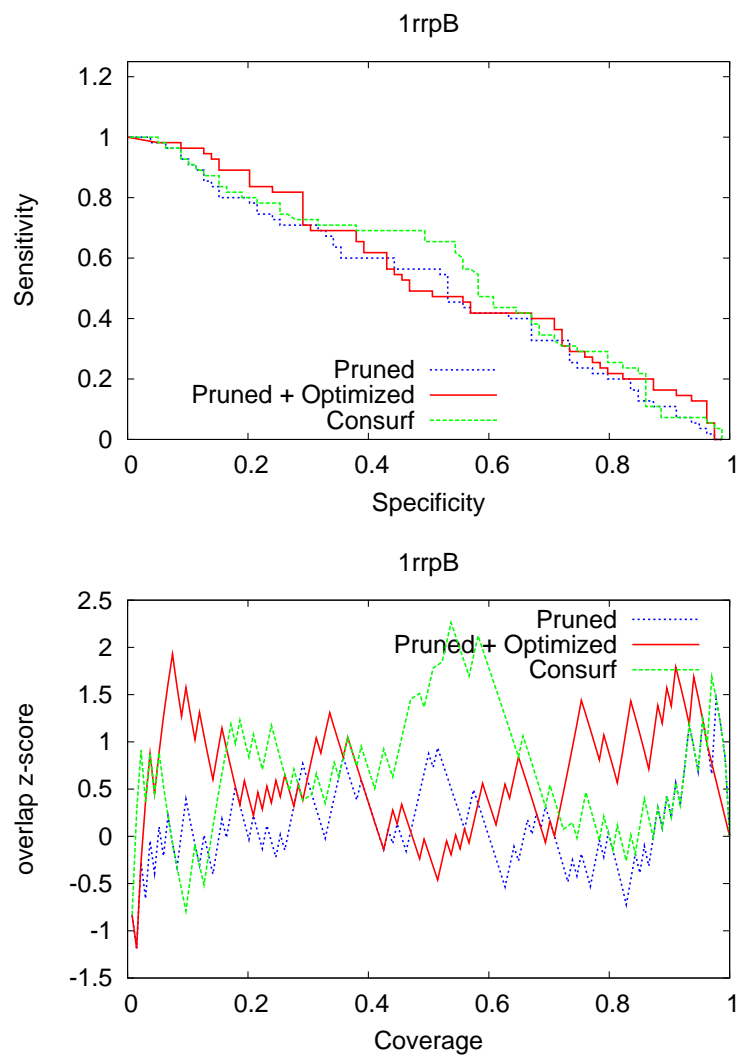
Figure 59:   a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1k7vA. b) The overlap z-scores as a function of coverage for the three methods is shown.
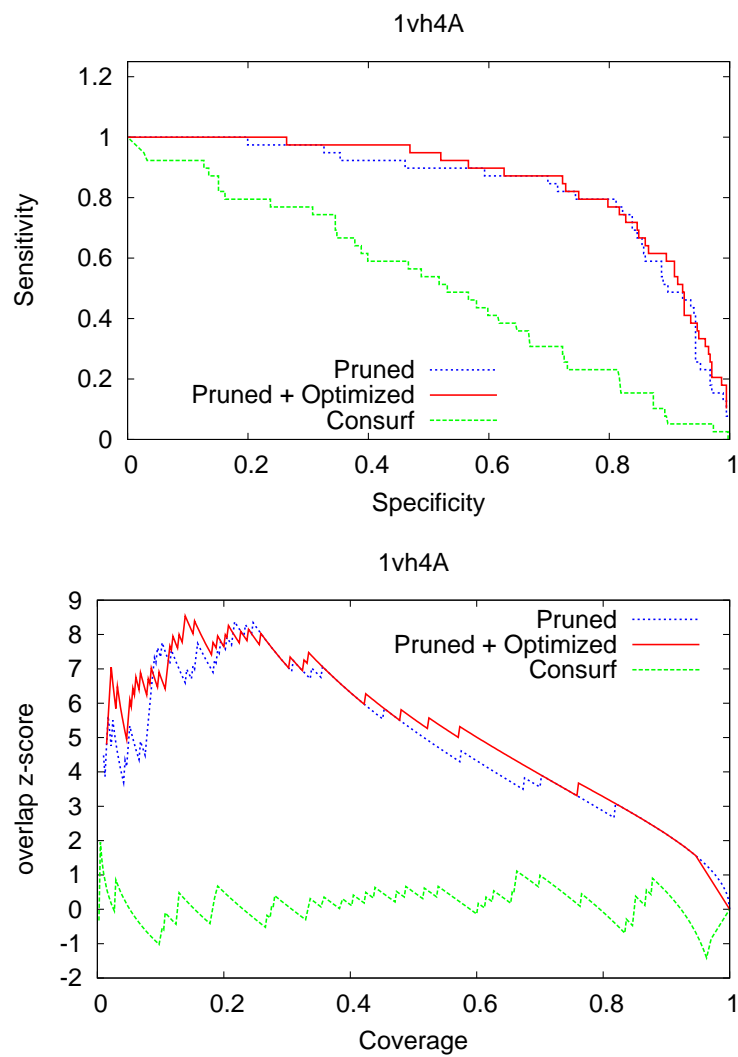
Figure 60: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ng1. b) The overlap z-scores as a function of coverage for the three methods is shown.
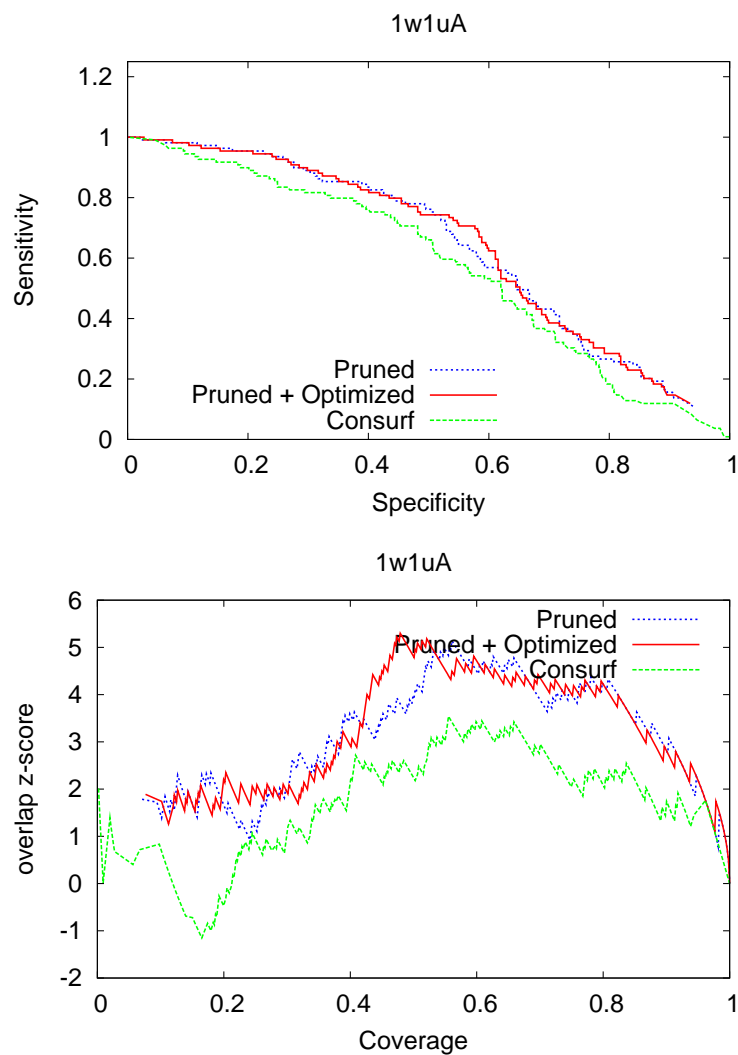
Figure 61: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1nzcA. b) The overlap z-scores as a function of coverage for the three methods is shown.
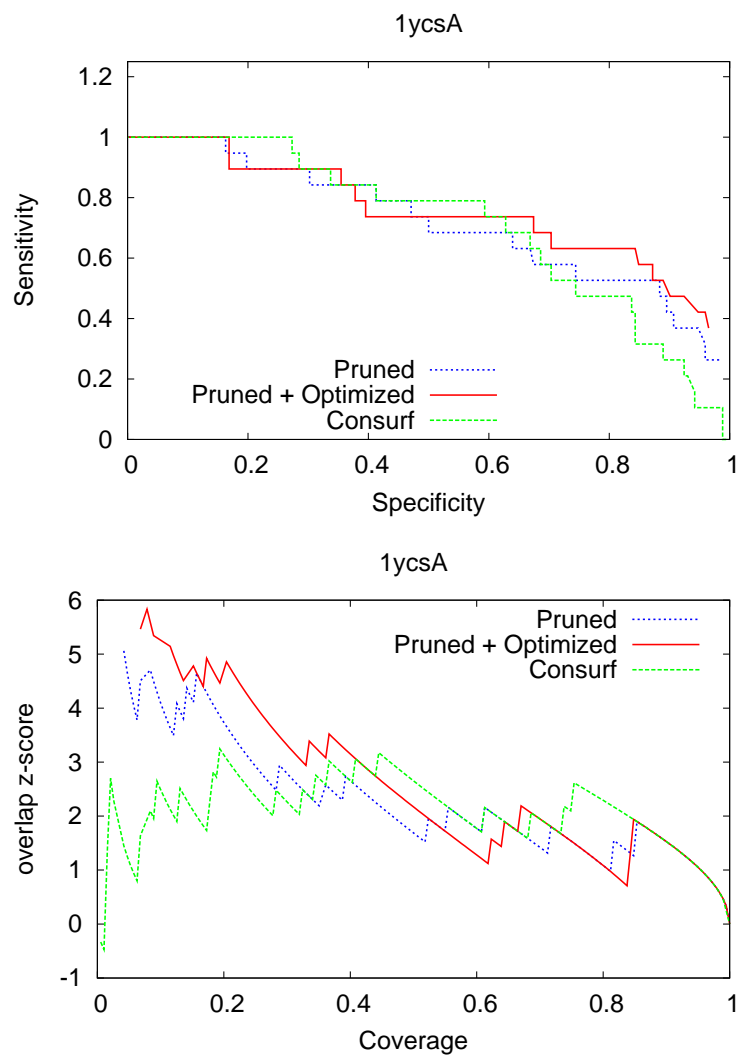
Figure 62:   a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pvdA. b) The overlap z-scores as a function of coverage for the three methods is shown.
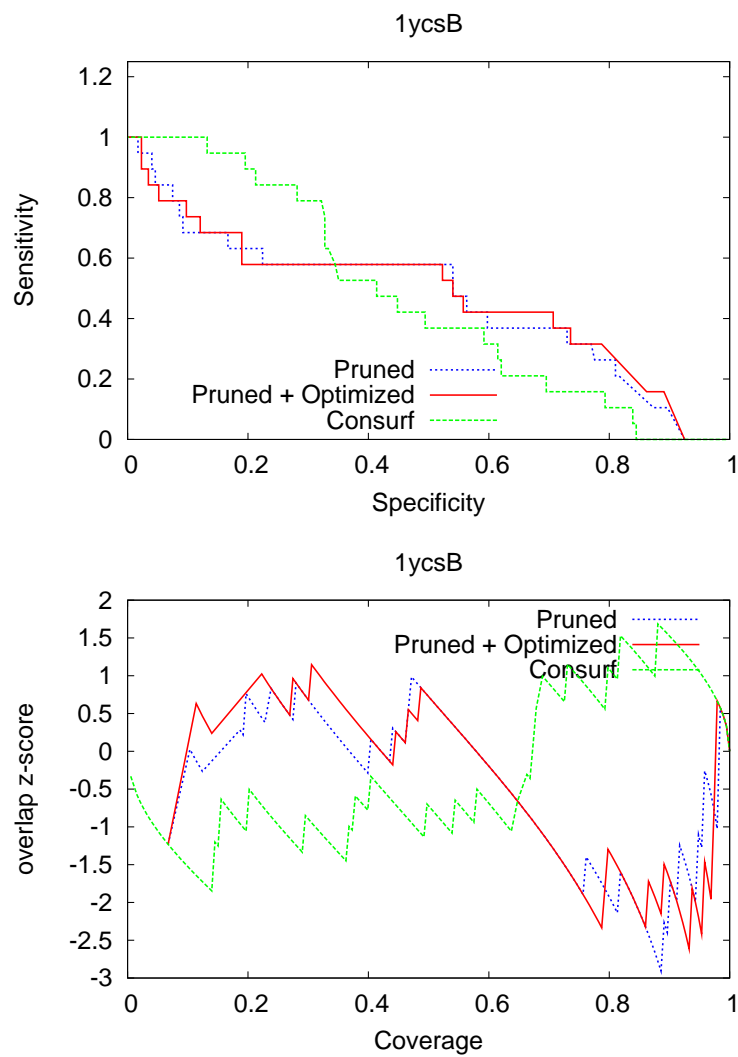
Figure 63: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1qumA. b) The overlap z-scores as a function of coverage for the three methods is shown.
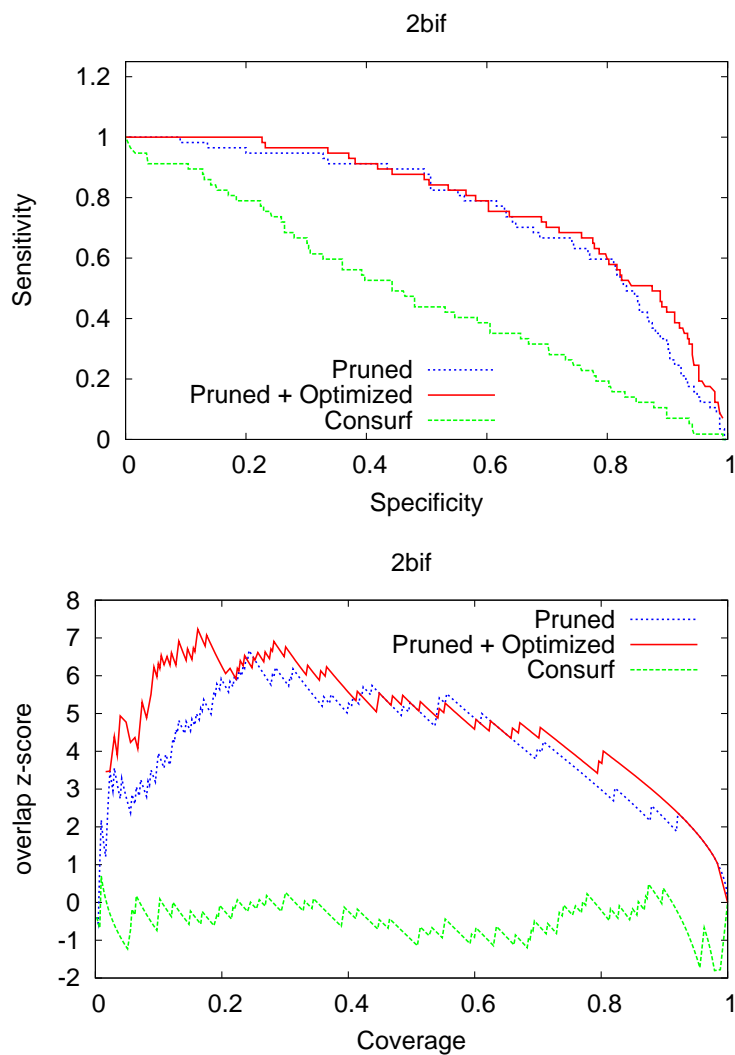
Figure 64: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1qupA. b) The overlap z-scores as a function of coverage for the three methods is shown.
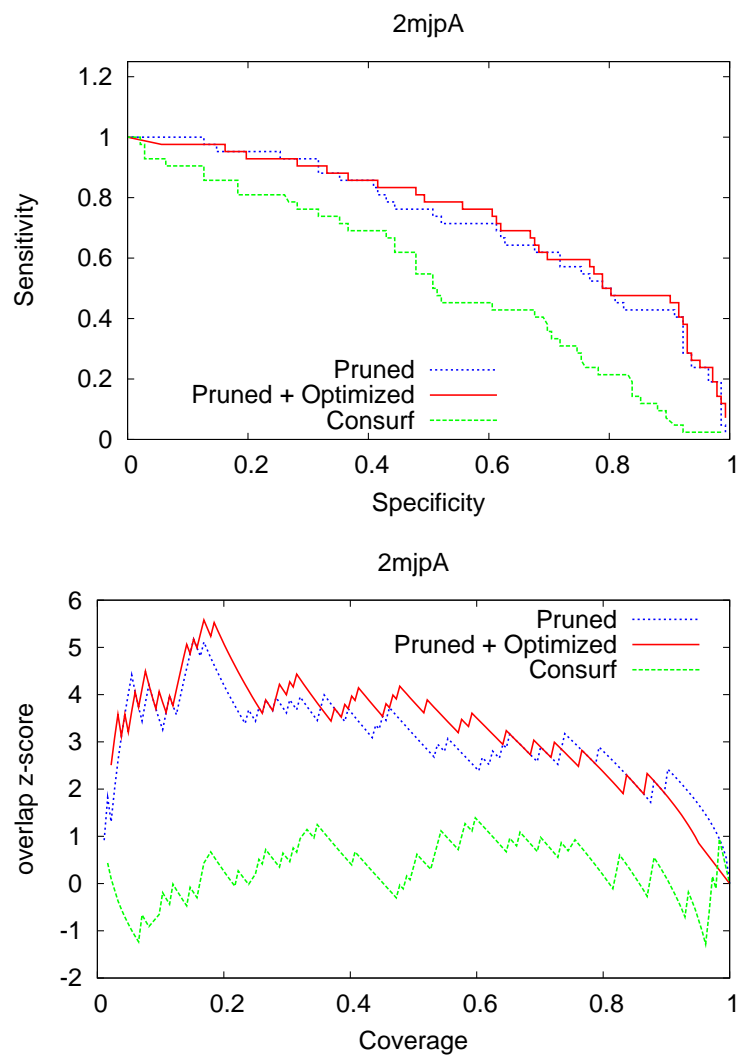
Figure 65: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1rrpA. b) The overlap z-scores as a function of coverage for the three methods is shown.
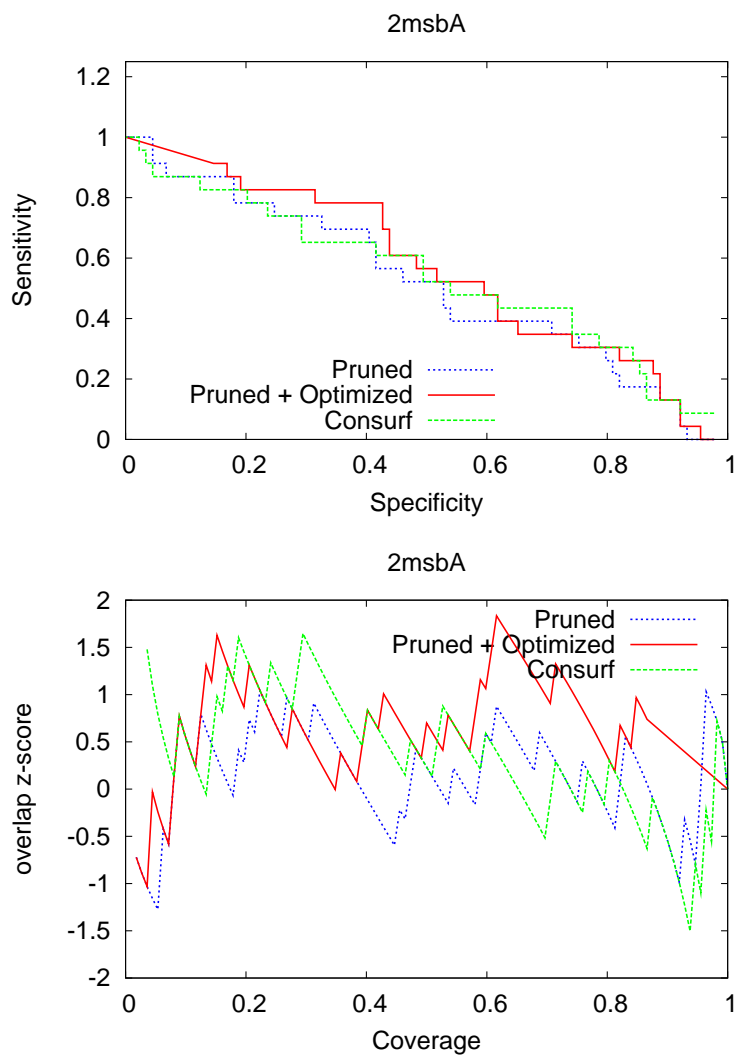
Figure 66: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1rrpB. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 67: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1vh4A. b) The overlap z-scores as a function of coverage for the three methods is shown.
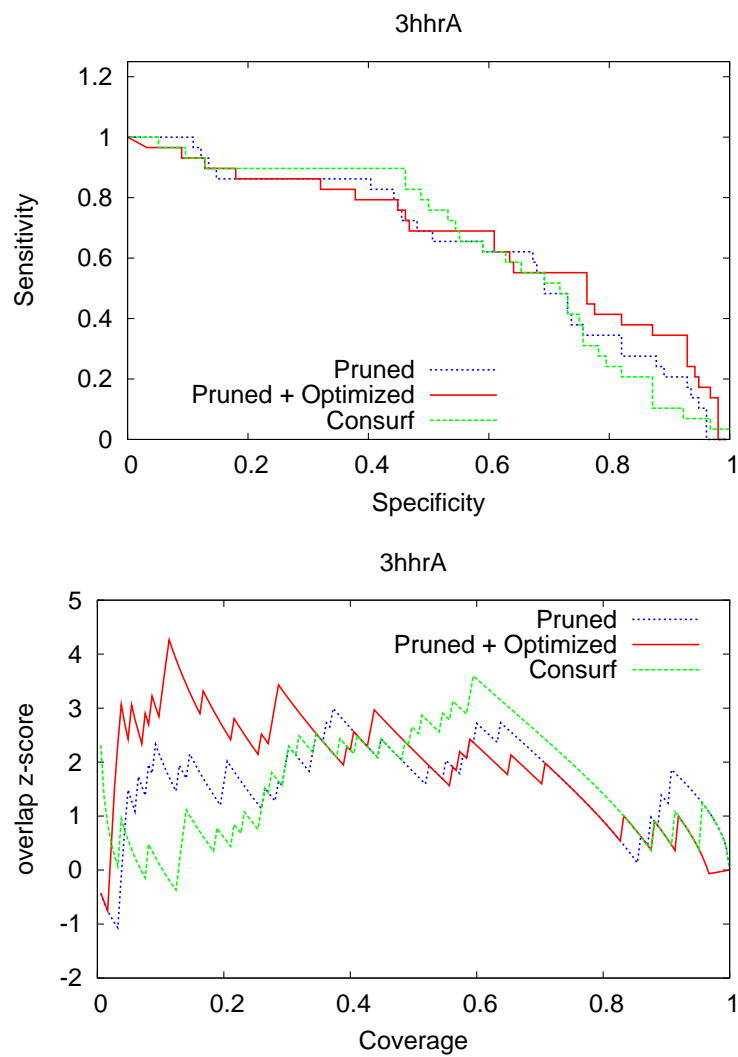
Figure 68: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1w1uA. b) The overlap z-scores as a function of coverage for the three methods is shown.
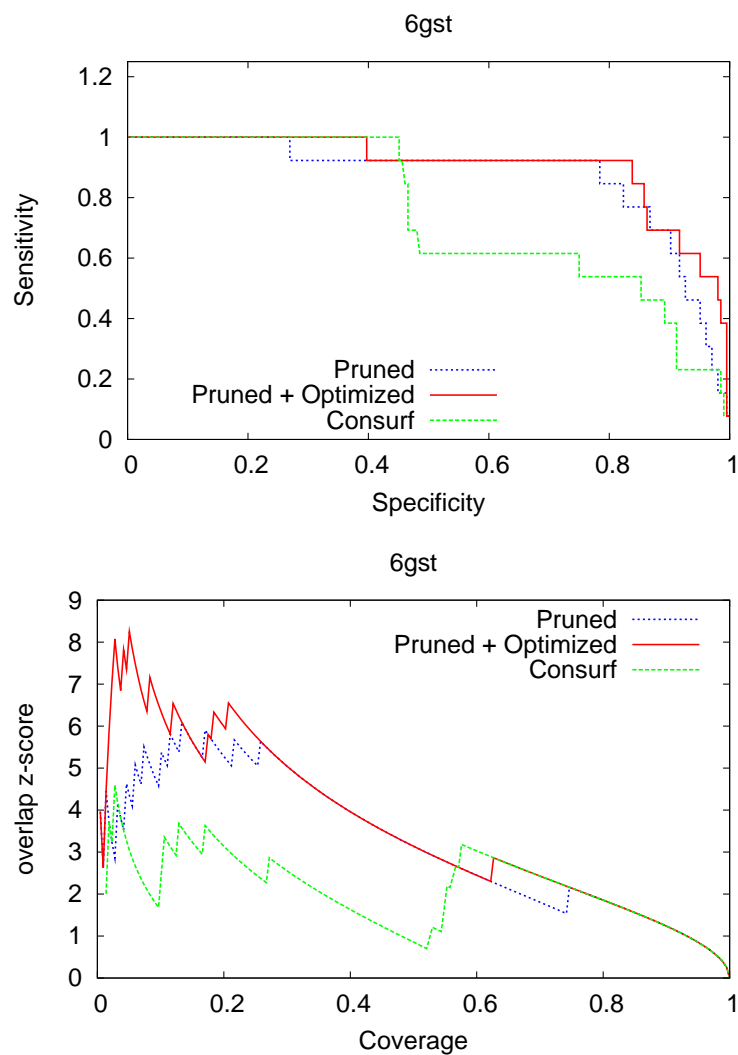
Figure 69: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ycsA. b) The overlap z-scores as a function of coverage for the three methods is shown.
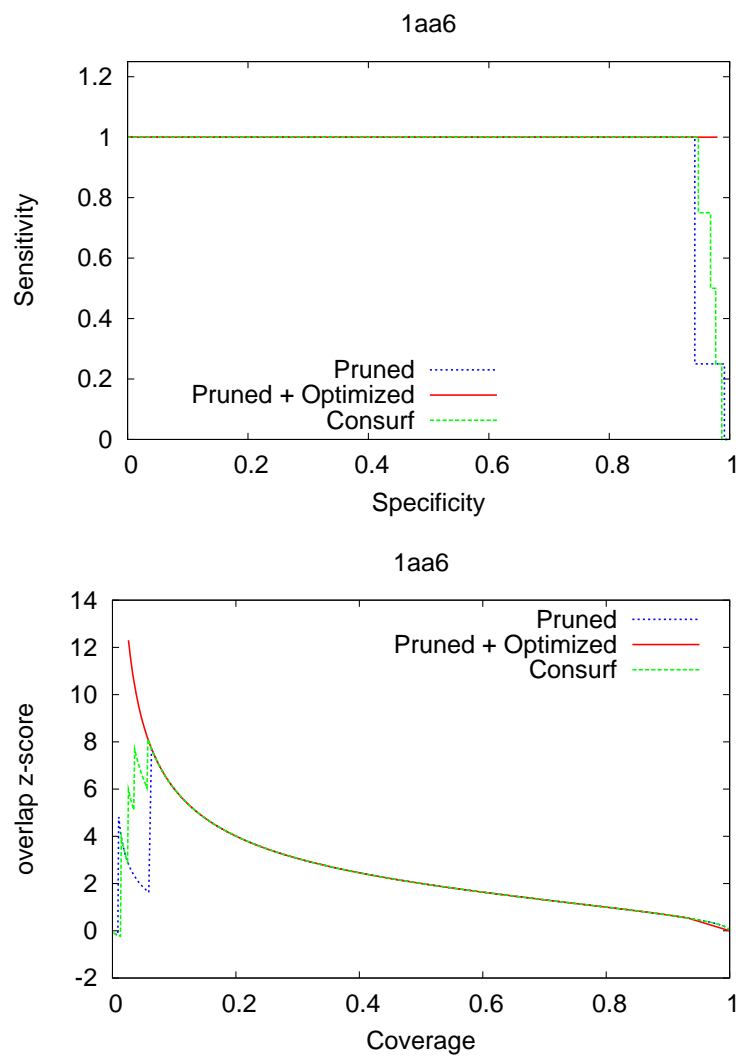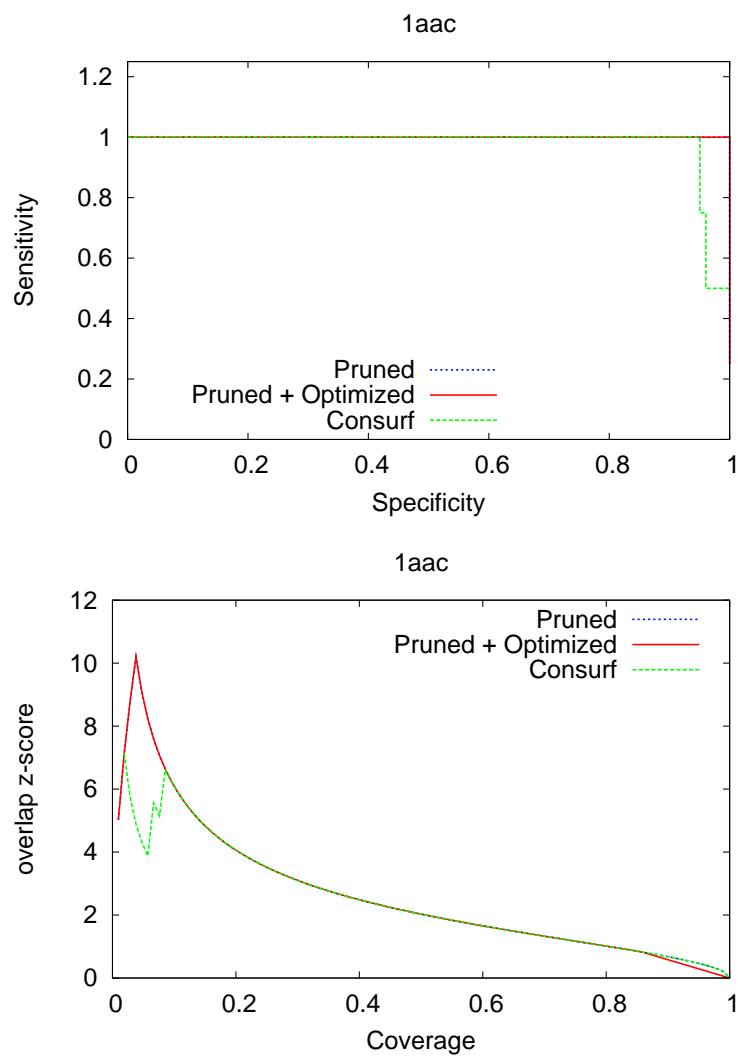
Figure 70: a) The sensitivity and specifity is shown for the pruned, pruned
+ optimized and the comparison method Consurf for PDBID 1ycsB. b) The
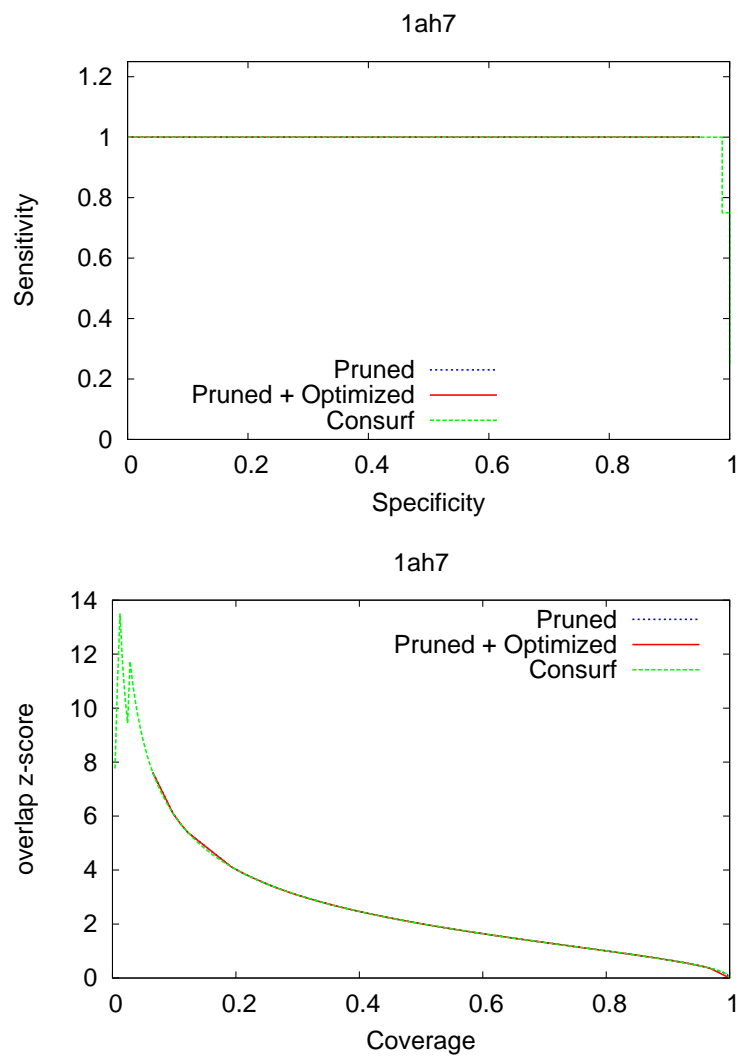overlap z-scores as a function of coverage for the three methods is shown.

Figure 71: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2bif. b) The overlap z-scores as a function of coverage for the three methods is shown.
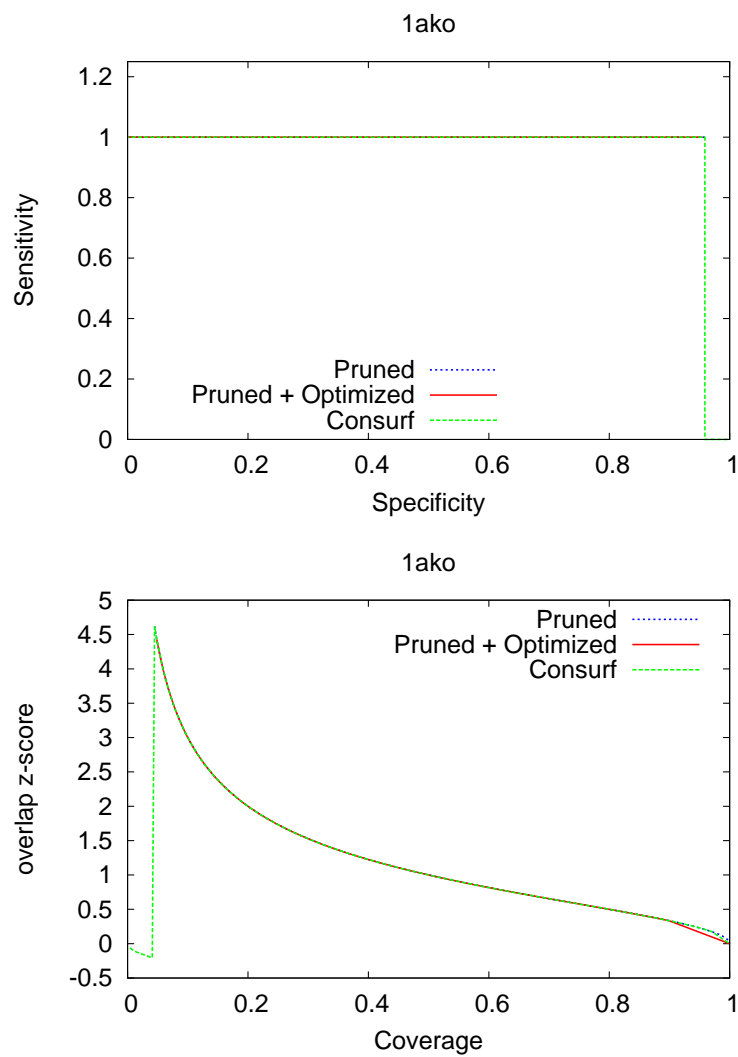
Figure 72: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2mjpA. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 73: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2msbA. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 74: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 3hhrA. b) The overlap z-scores as a function of coverage for the three methods is shown.
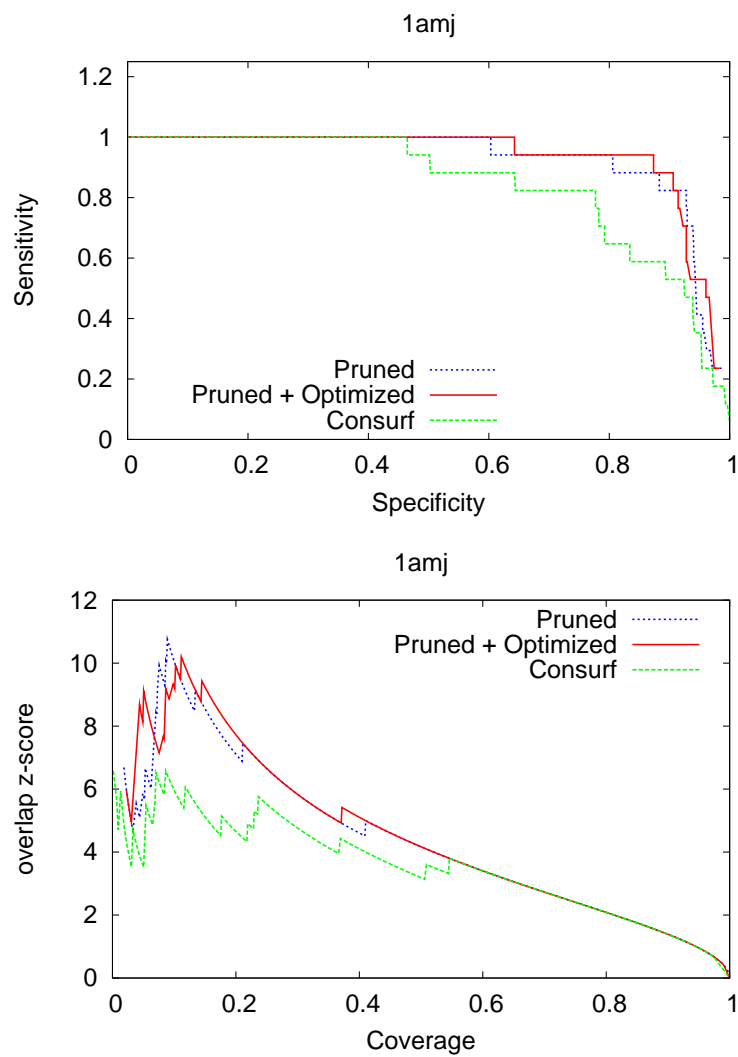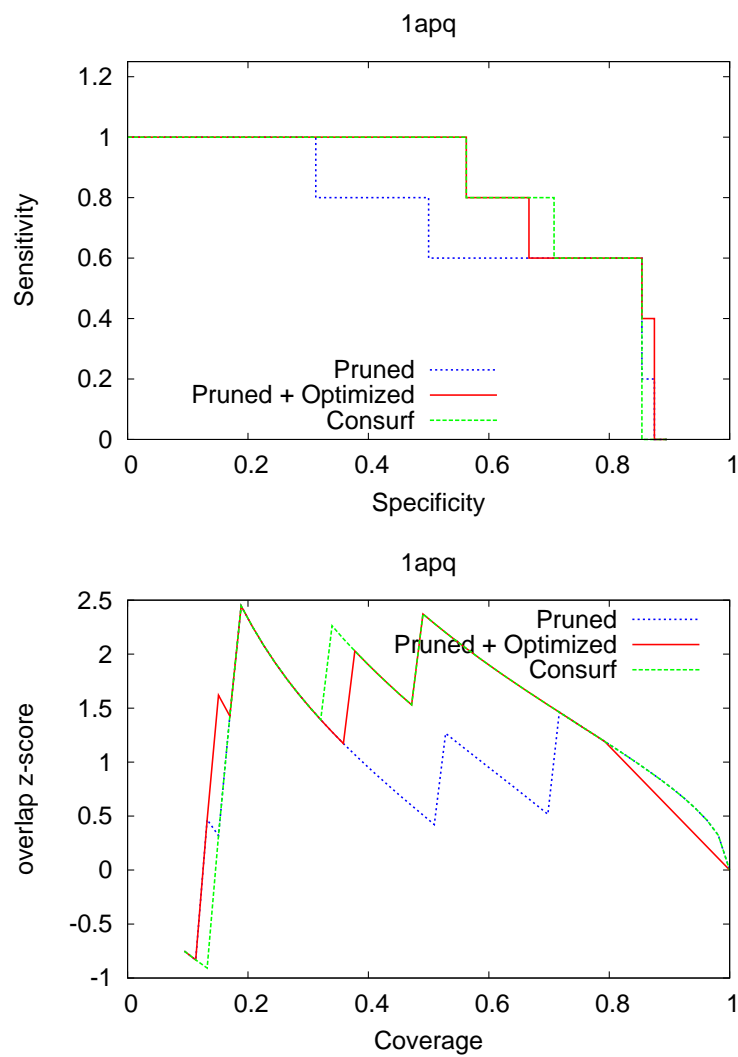
Figure 75: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 6gst. b) The overlap z-scores as a function of coverage for the three methods is shown.

Figure 76: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aa6. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
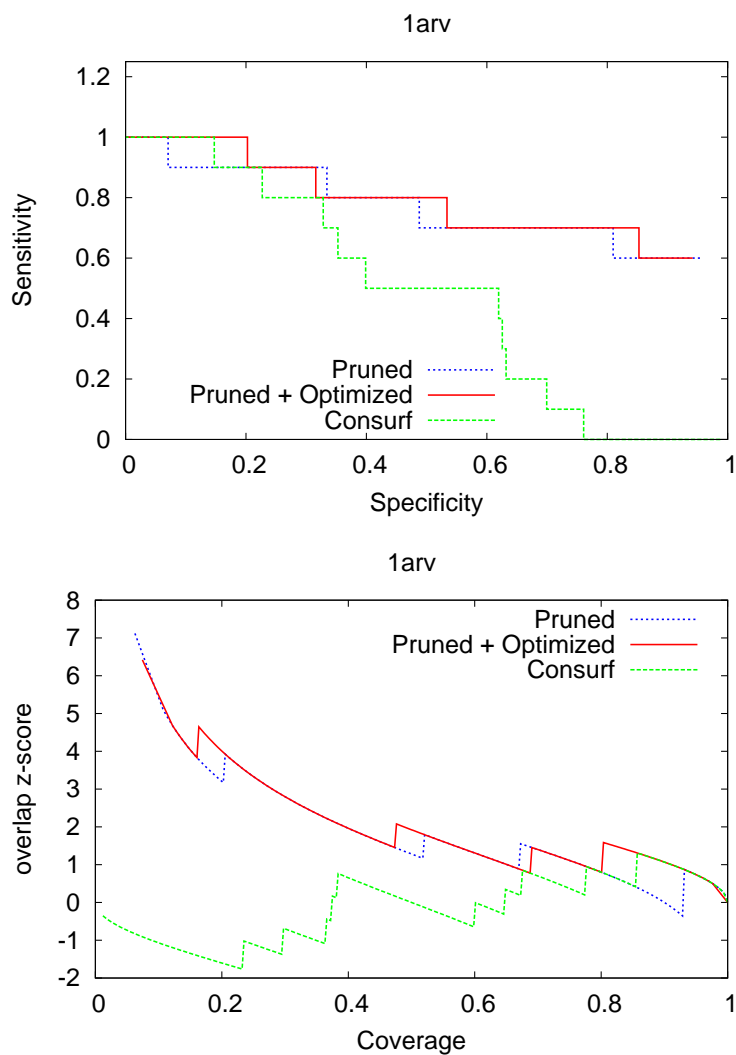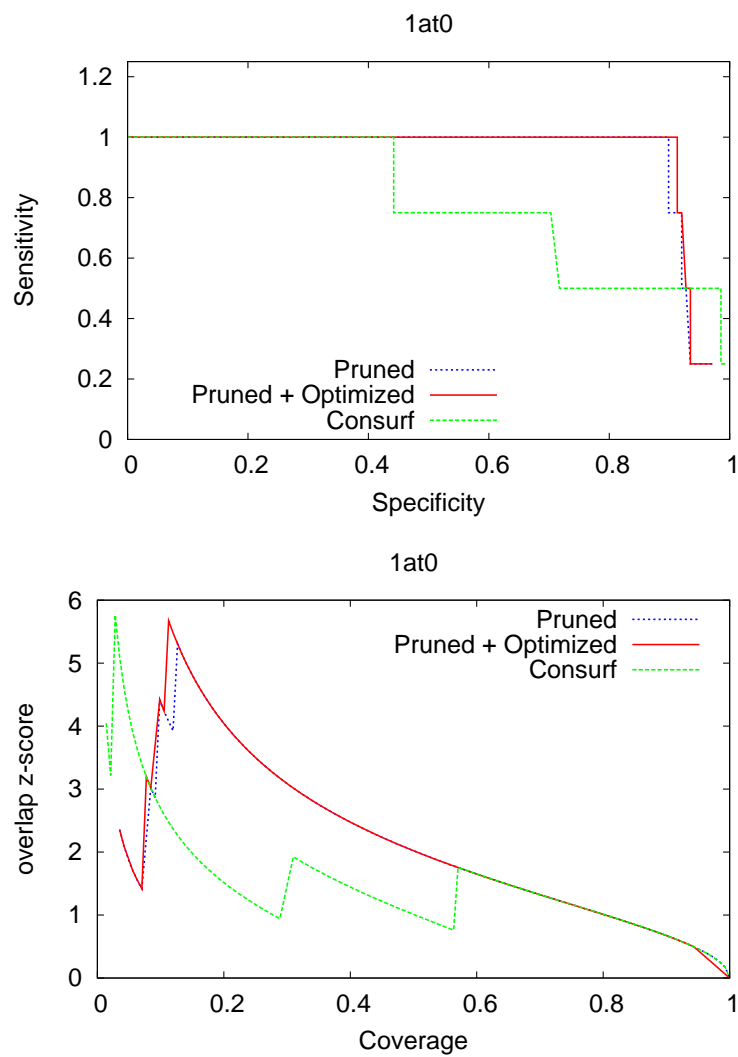
Figure 77: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1aac. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
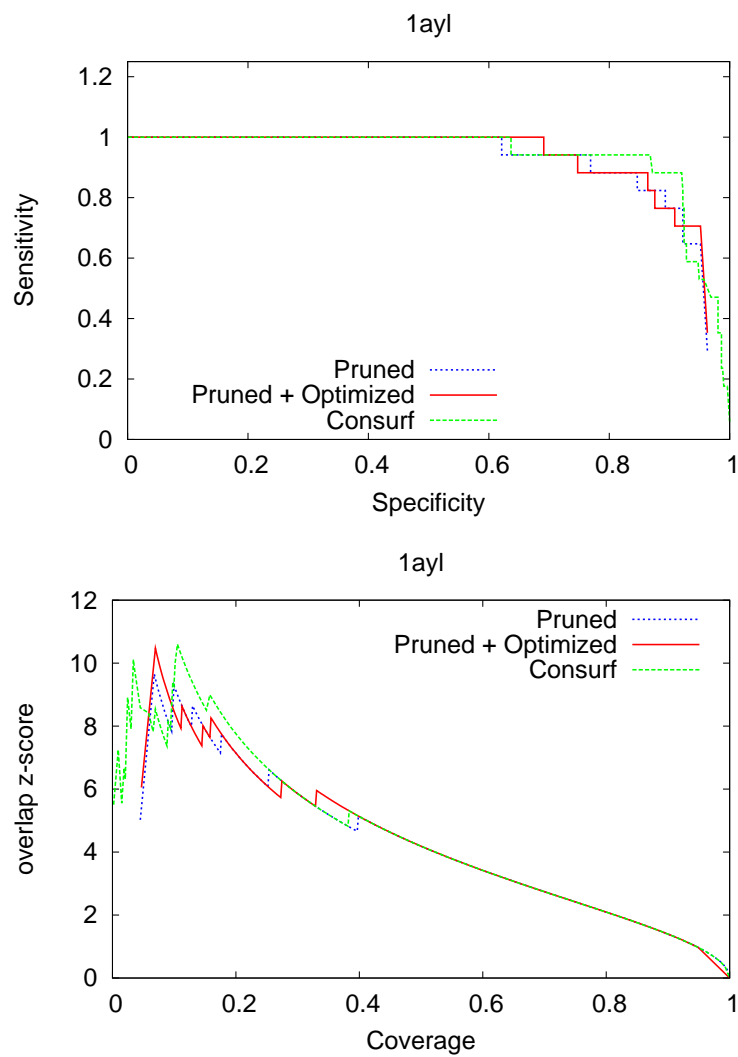
Figure 78: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ah7. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
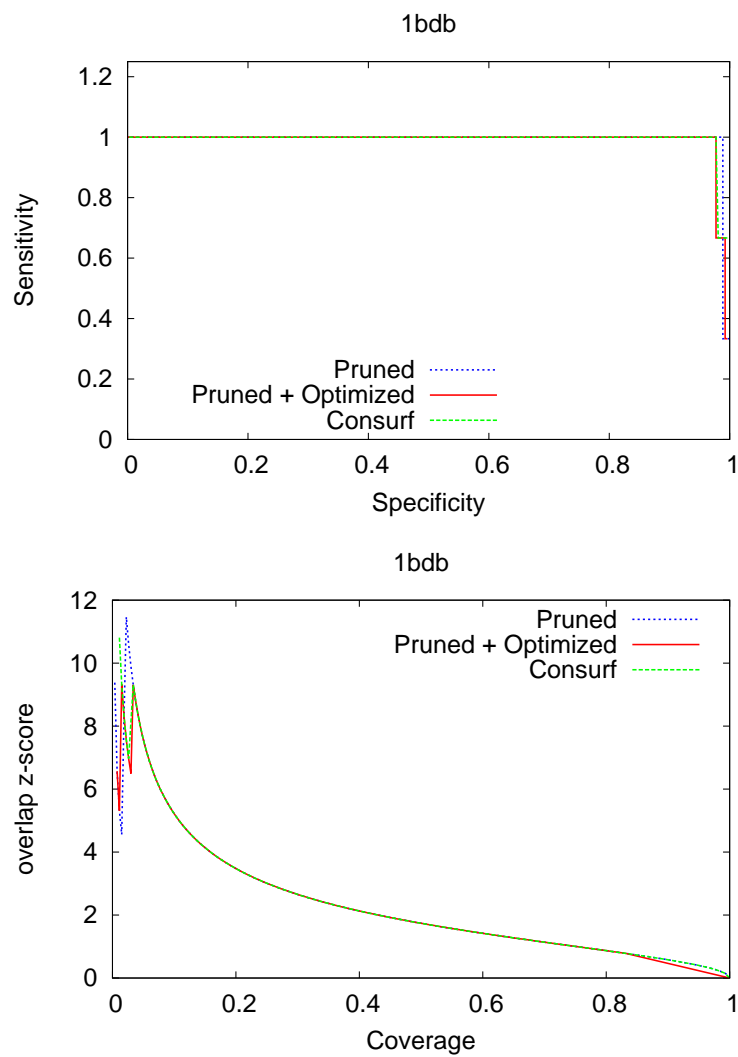
Figure 79: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ako. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
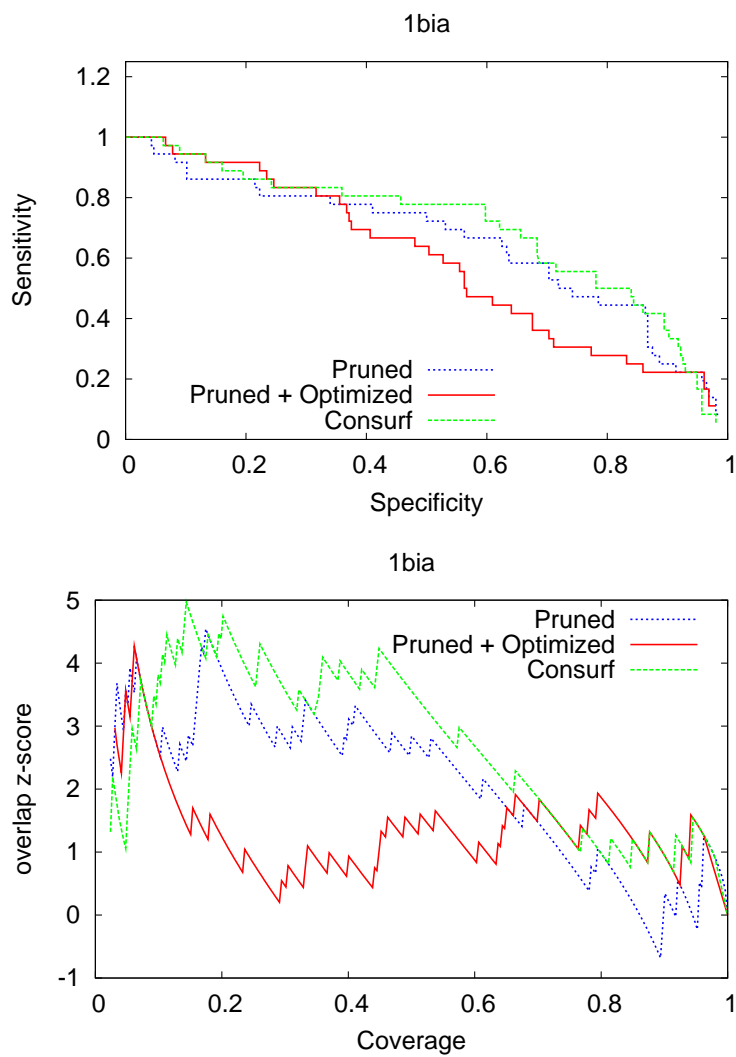
Figure 80: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1amj. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 81: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1apq. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
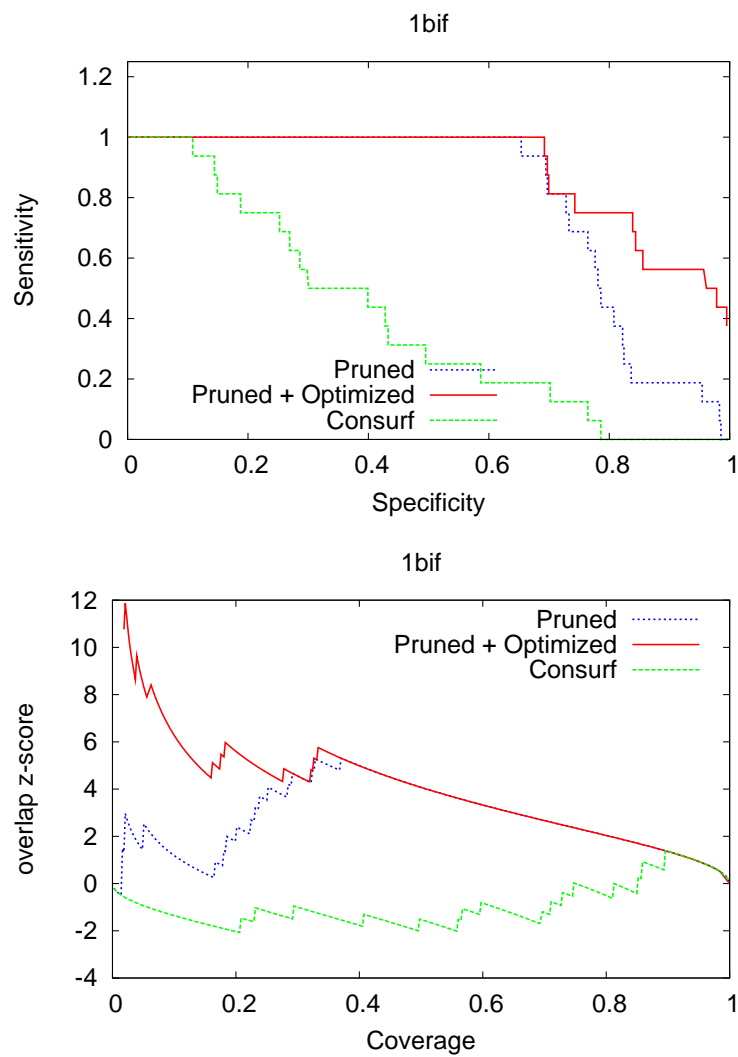
Figure 82: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1arv. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
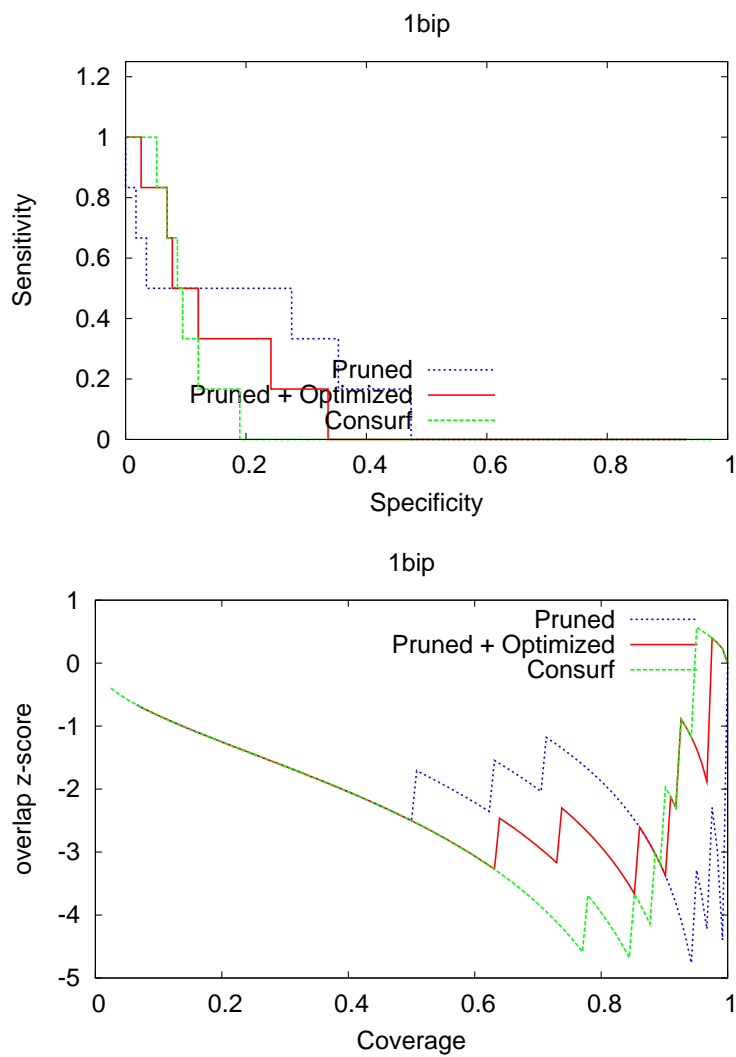
Figure 83: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1at0. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
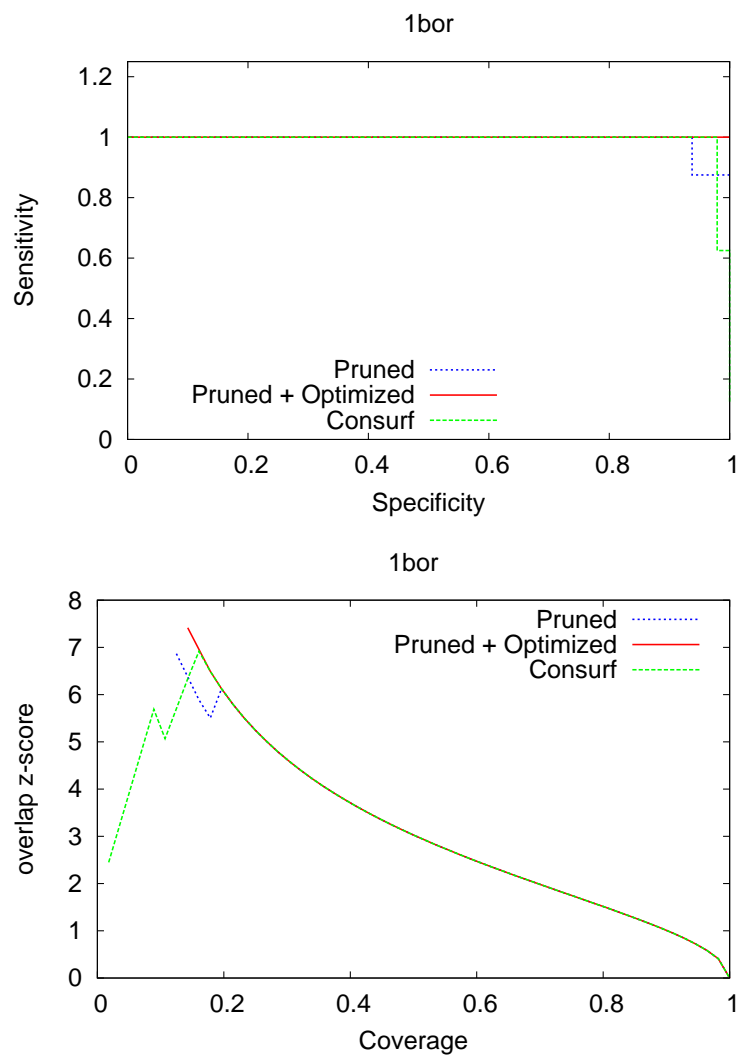
Figure 84: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ayl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
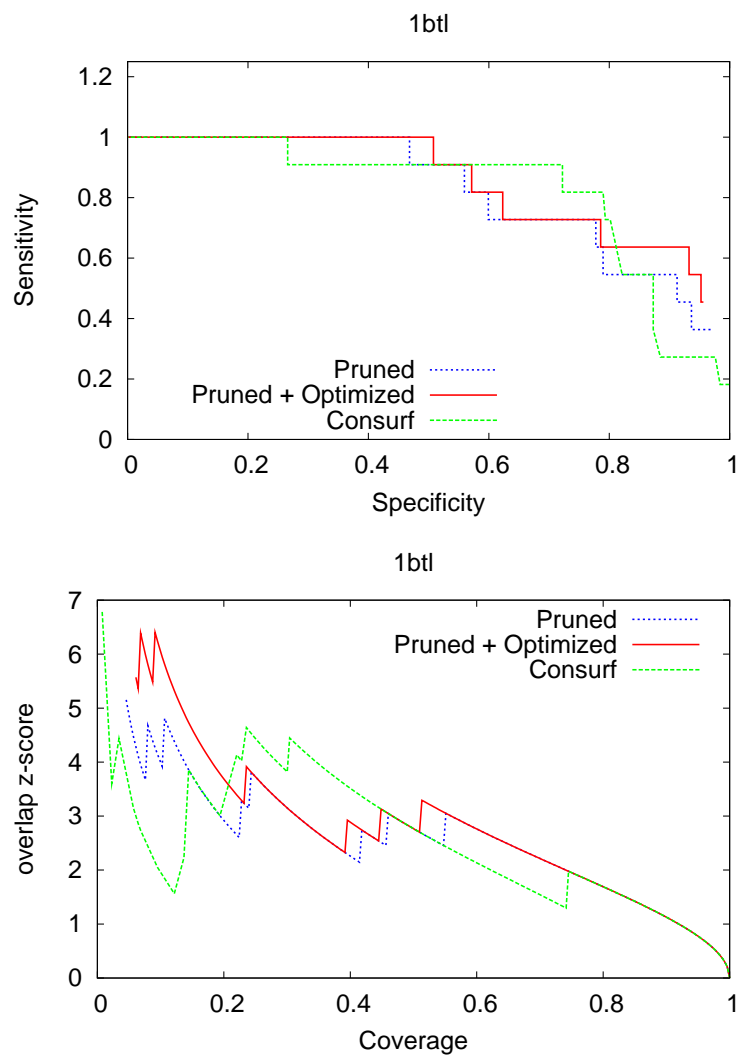
Figure 85: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bdb. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
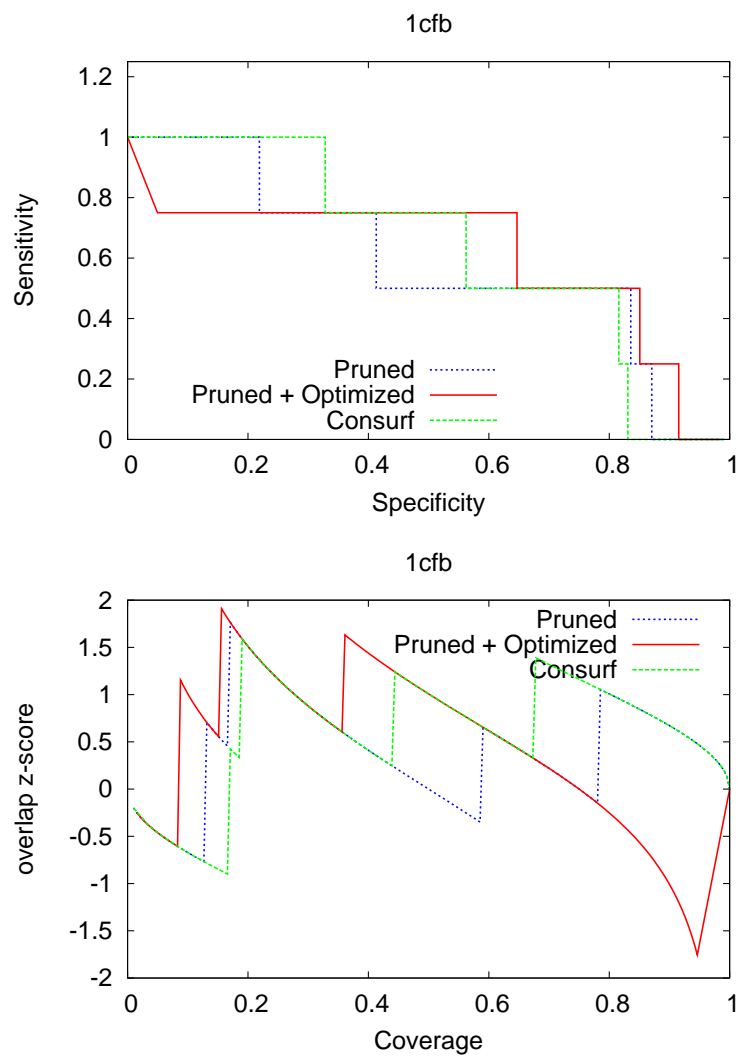
Figure 86: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bia. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
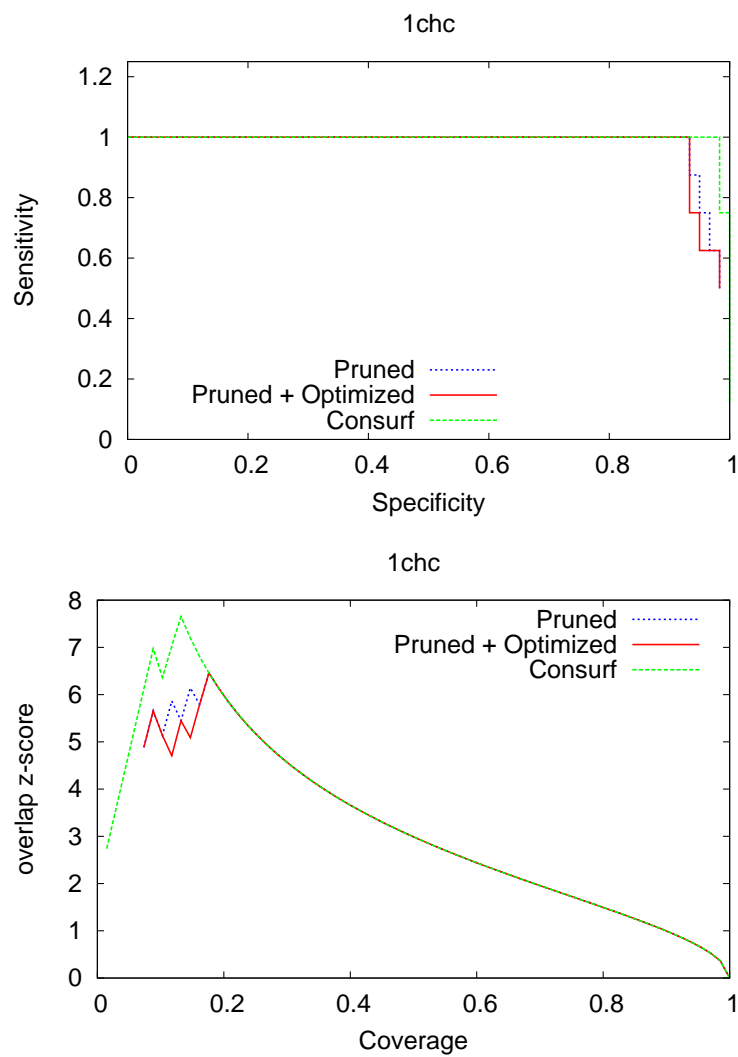
Figure 87: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bif. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
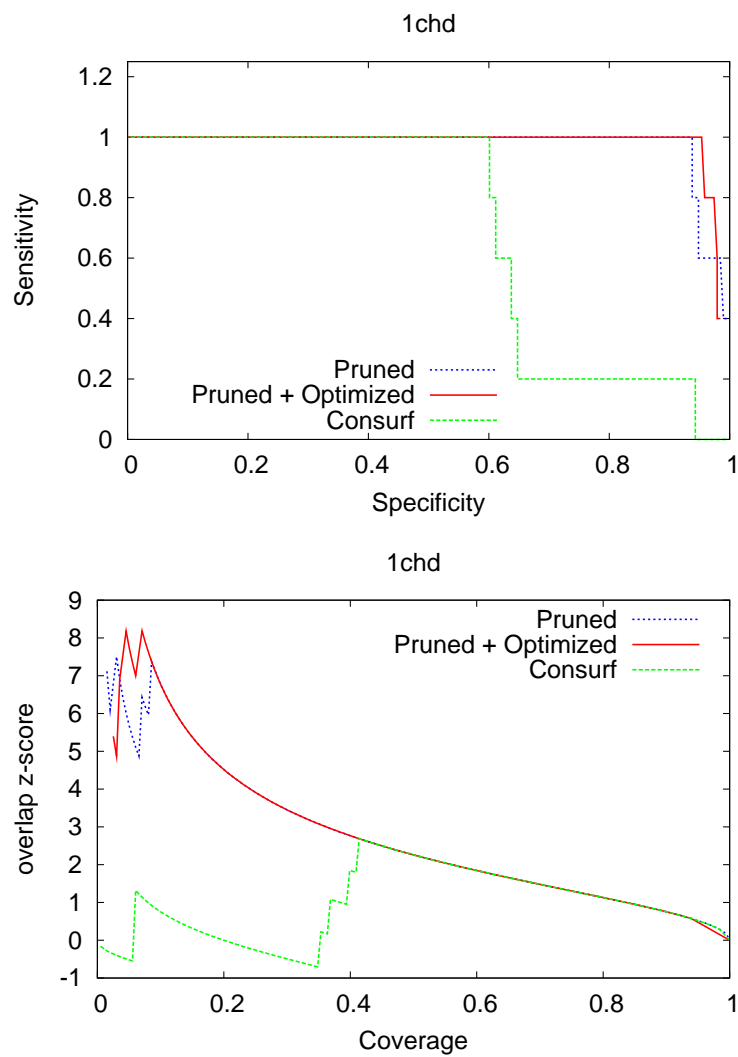
Figure 88: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bip. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
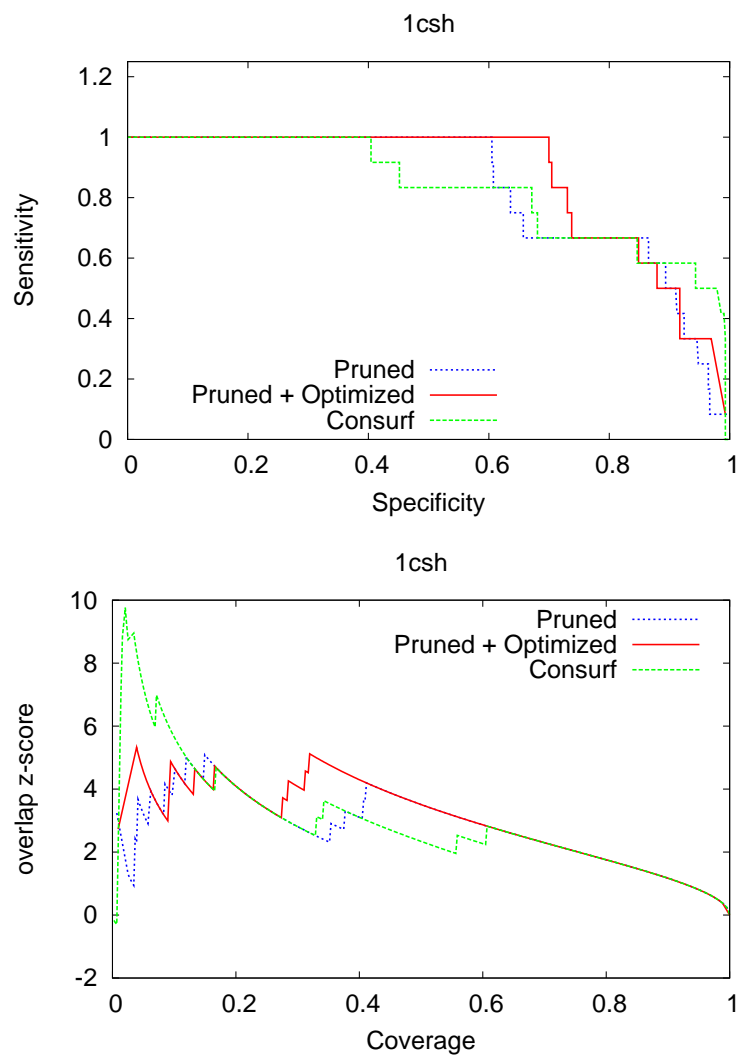
Figure 89: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1bor. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
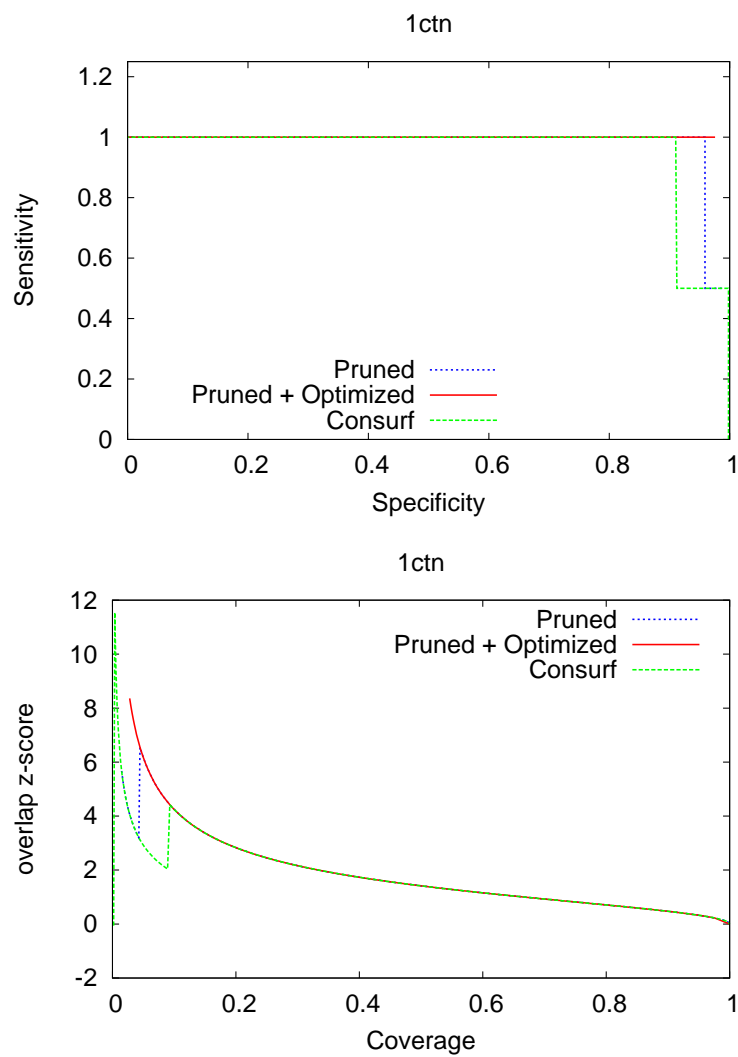
Figure 90: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1btl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
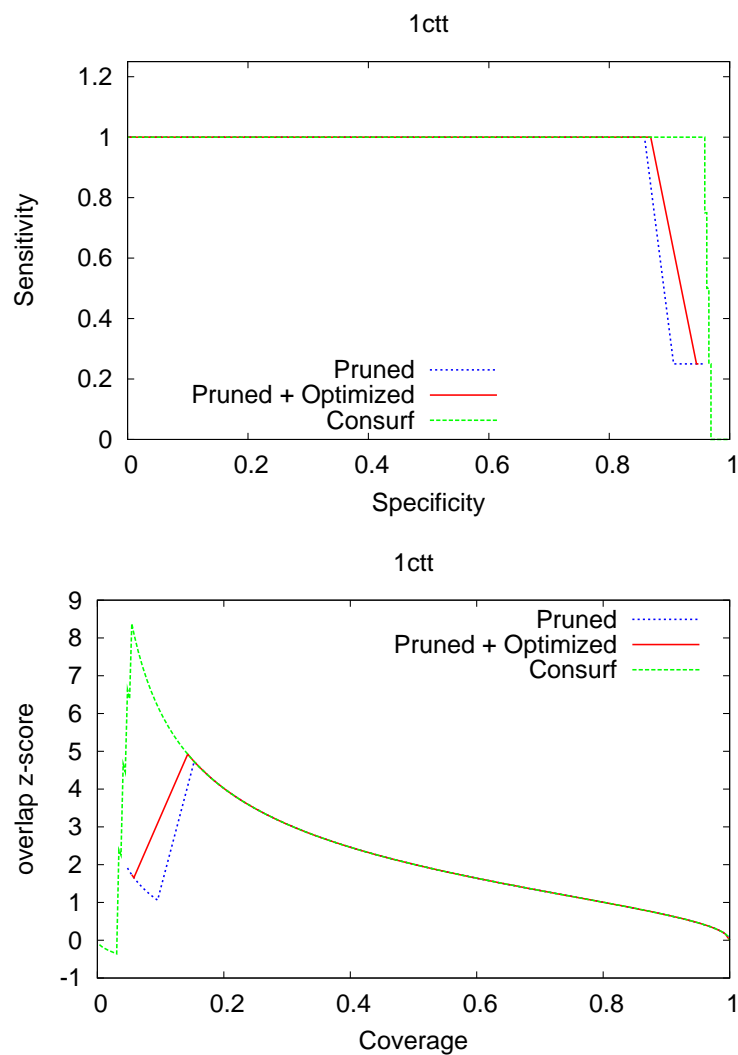
Figure 91: a) The sensitivity and specifity is shown for the pruned, pruned +
optimized and the comparison method Consurf for PDBID 1cfb. b) The overlap
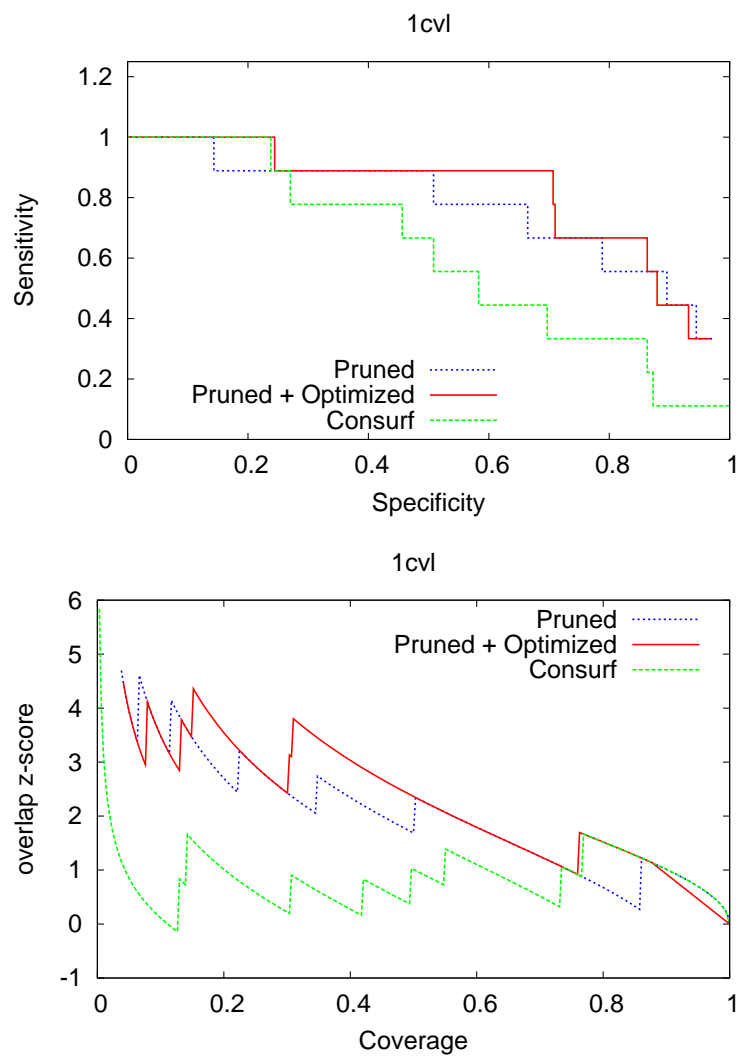z-scores as a function of rank coverage for the three methods is shown.

Figure 92: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1chc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
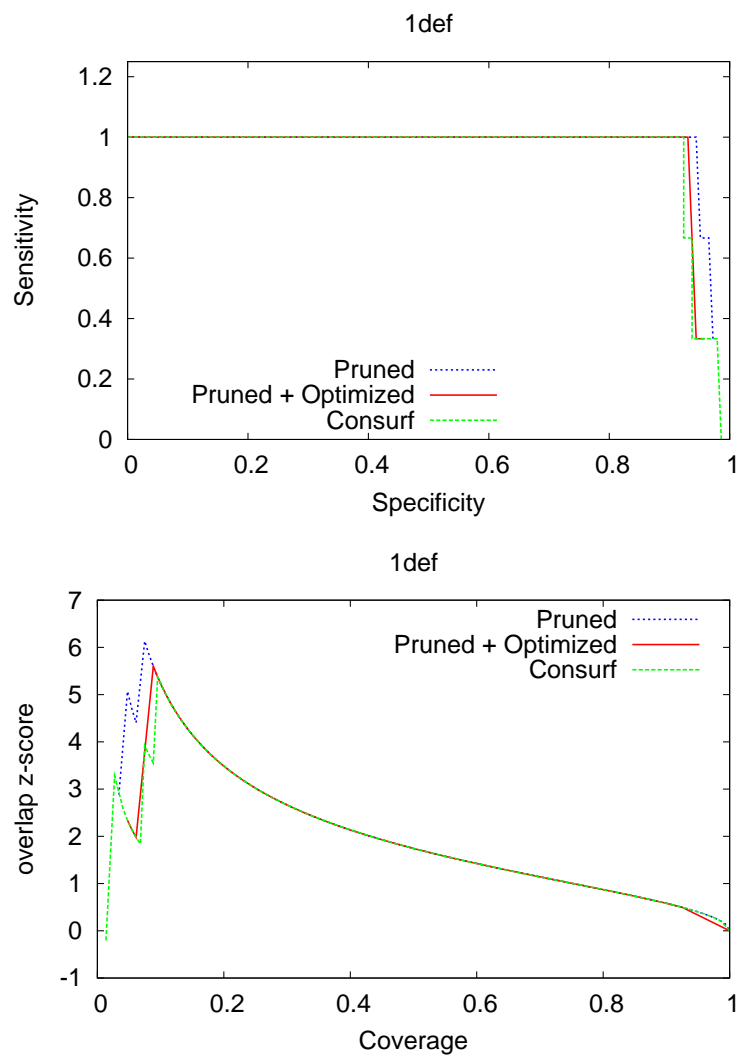
Figure 93: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1chd. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
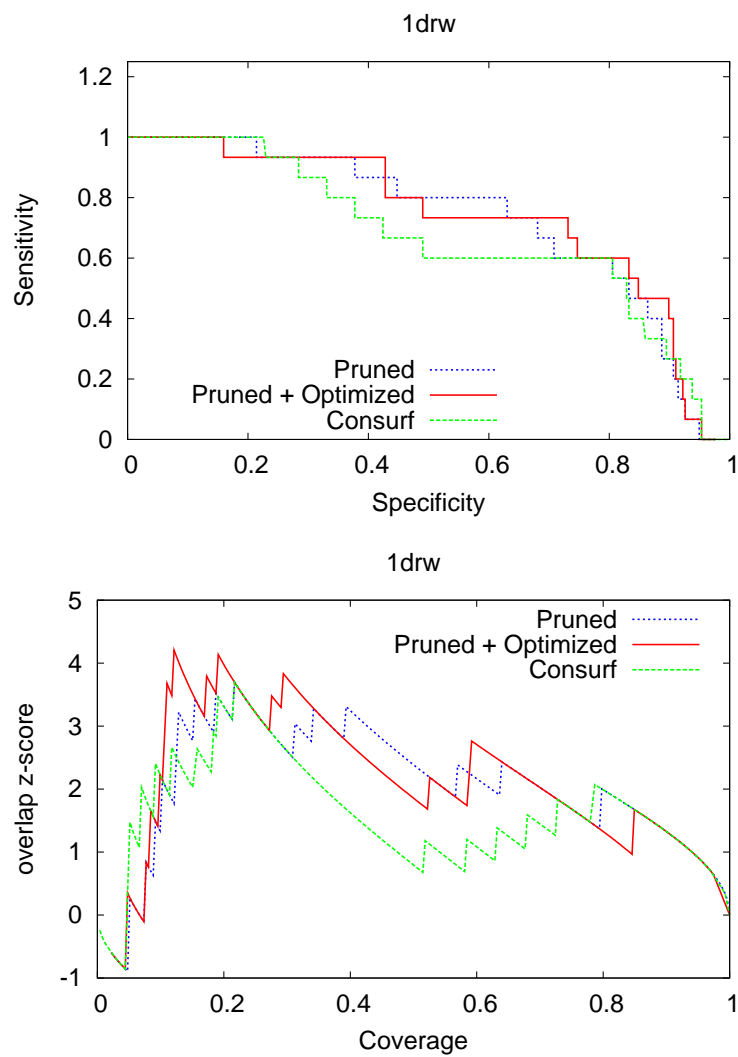
Figure 94: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1csh. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
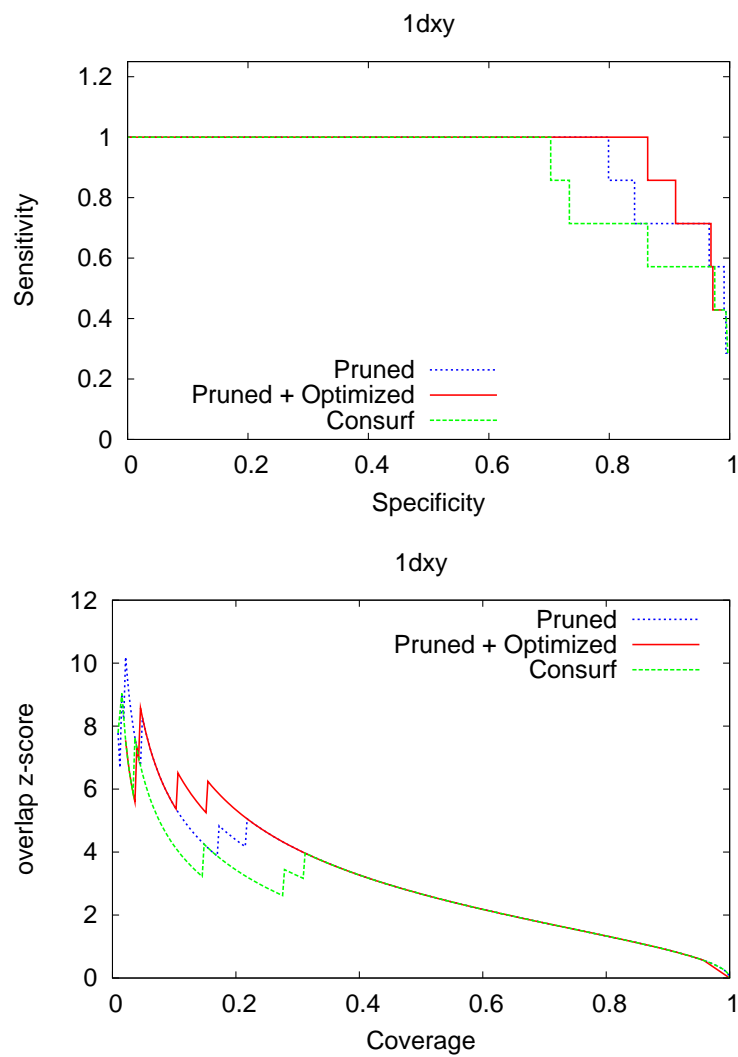
Figure 95: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ctn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
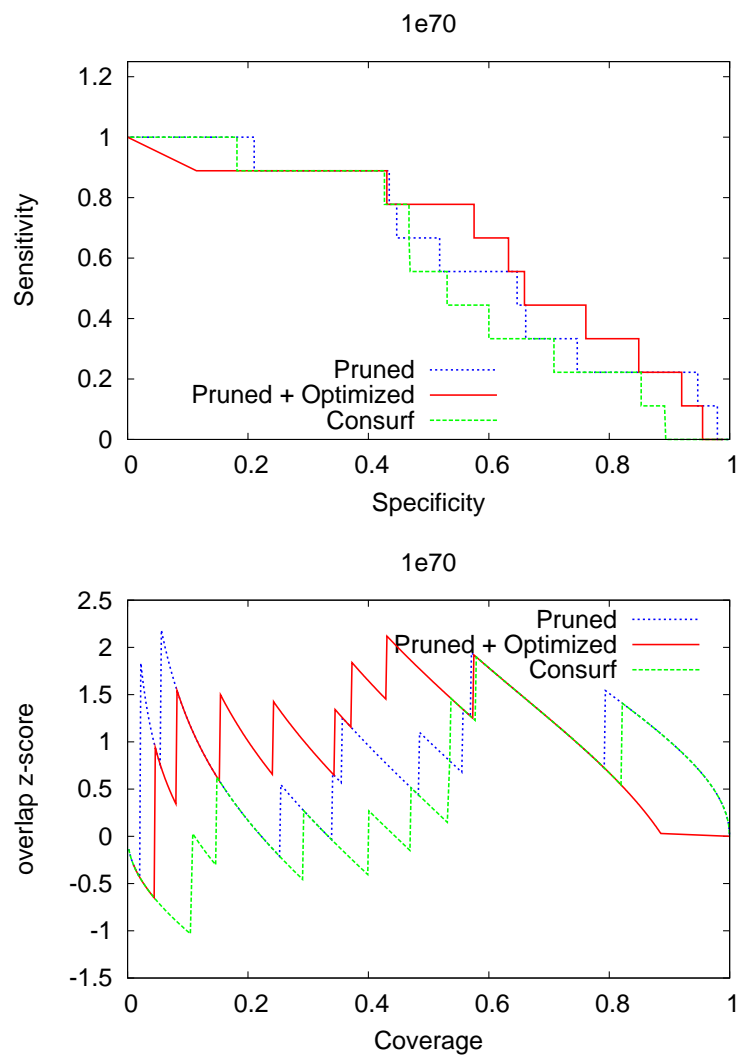
Figure 96: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ctt. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 97: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1cvl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
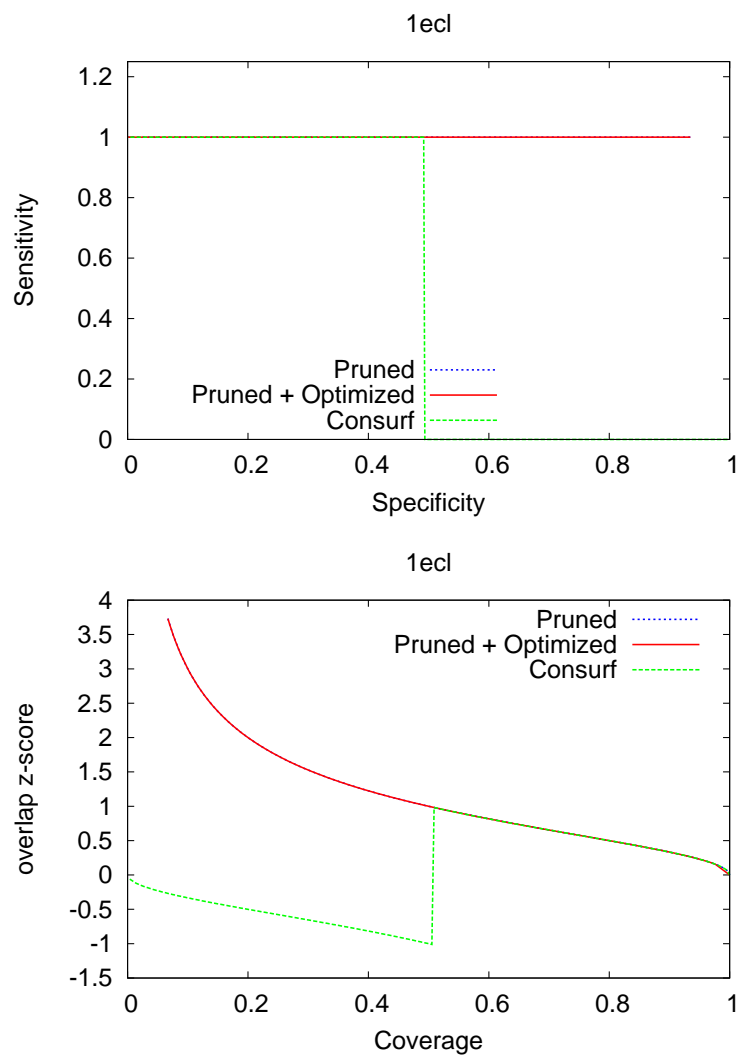
Figure 98: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1def. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
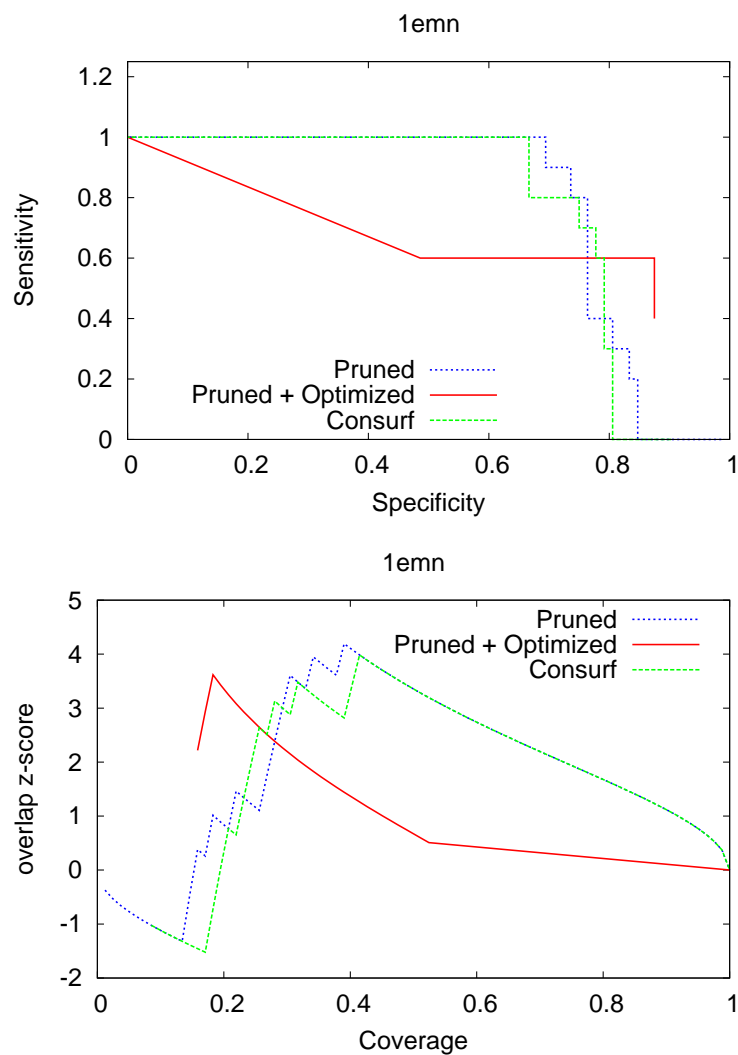
Figure 99: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1drw. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
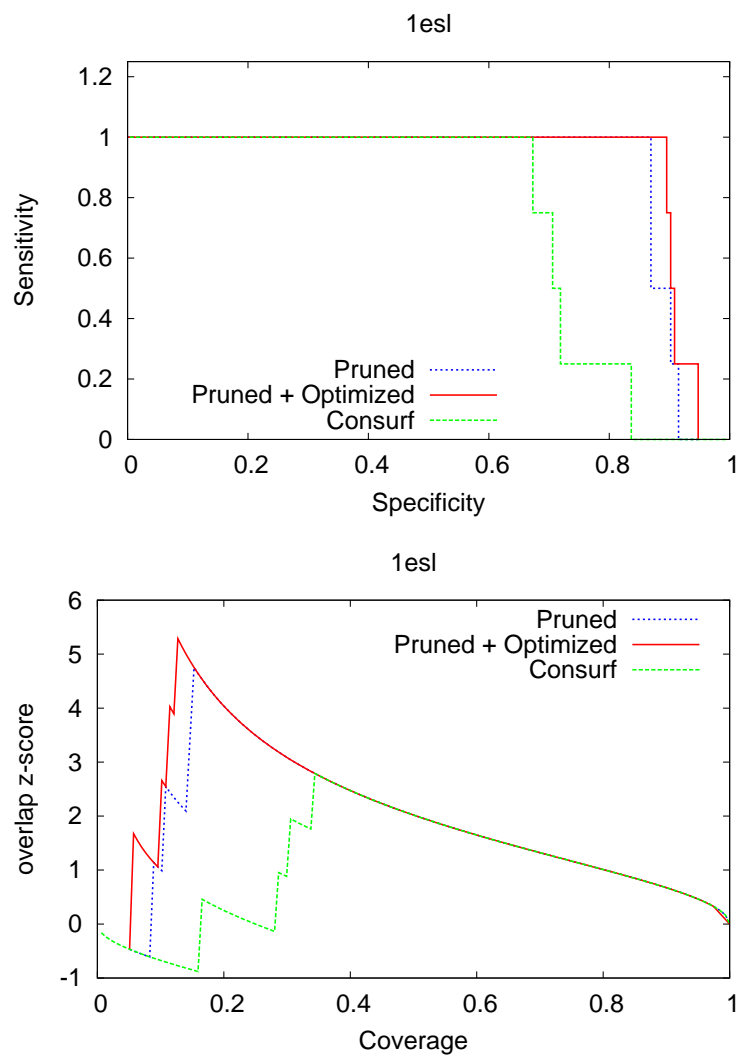
Figure 100: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1dxy. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
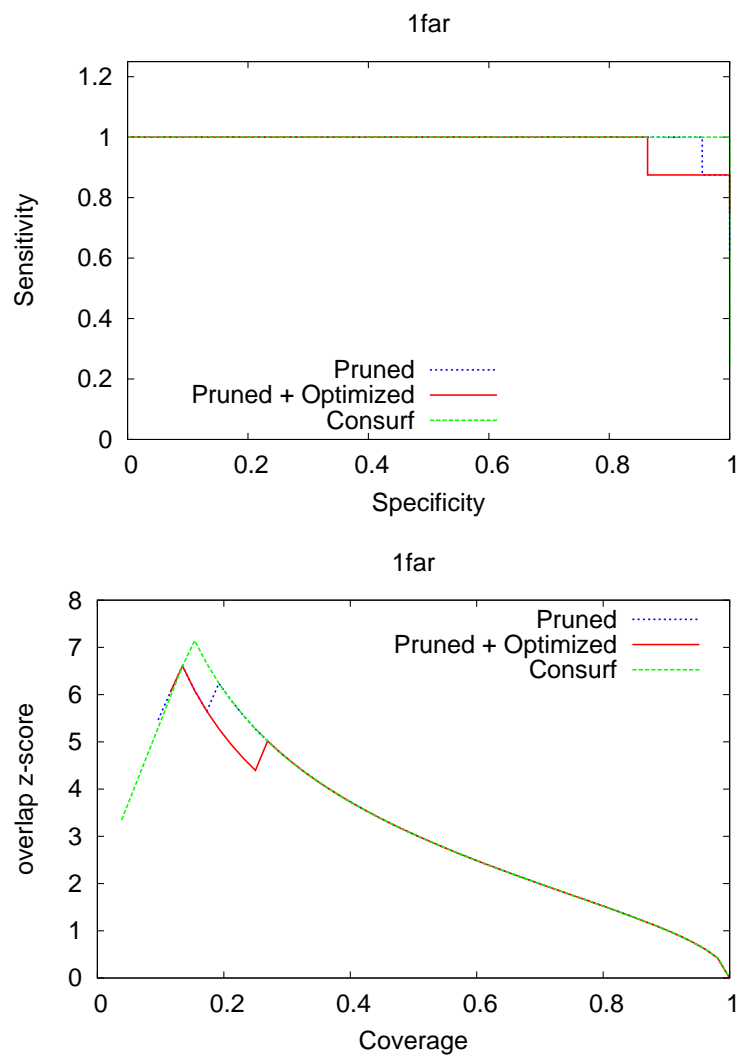
Figure 101: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1e70. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
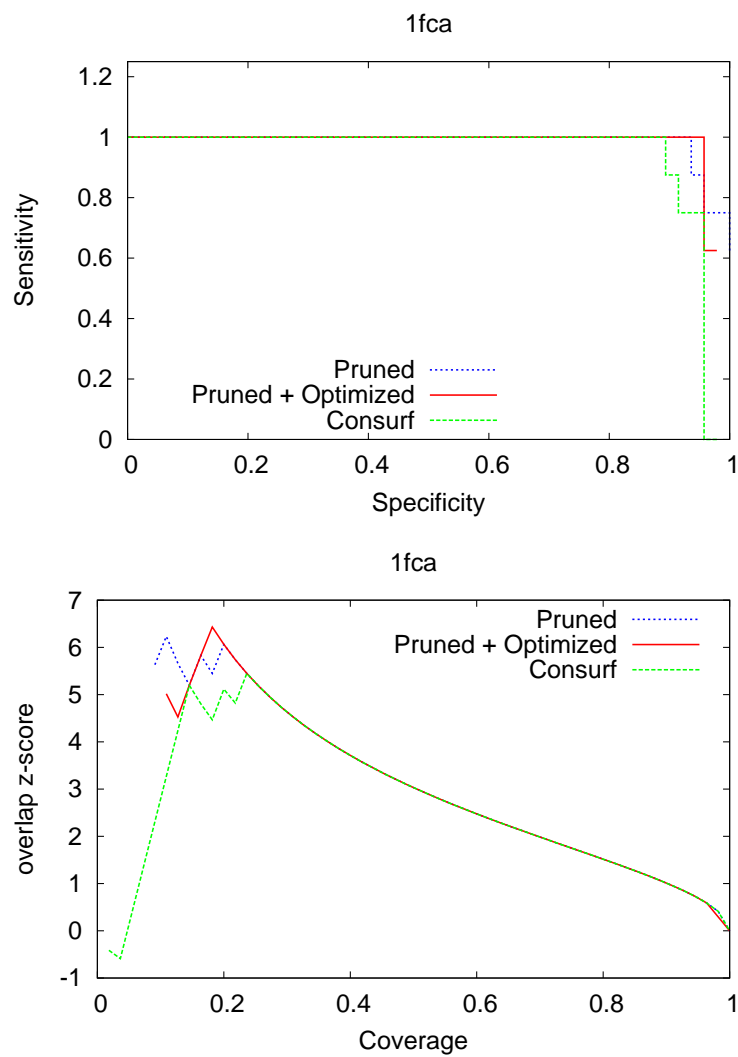
Figure 102: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ecl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
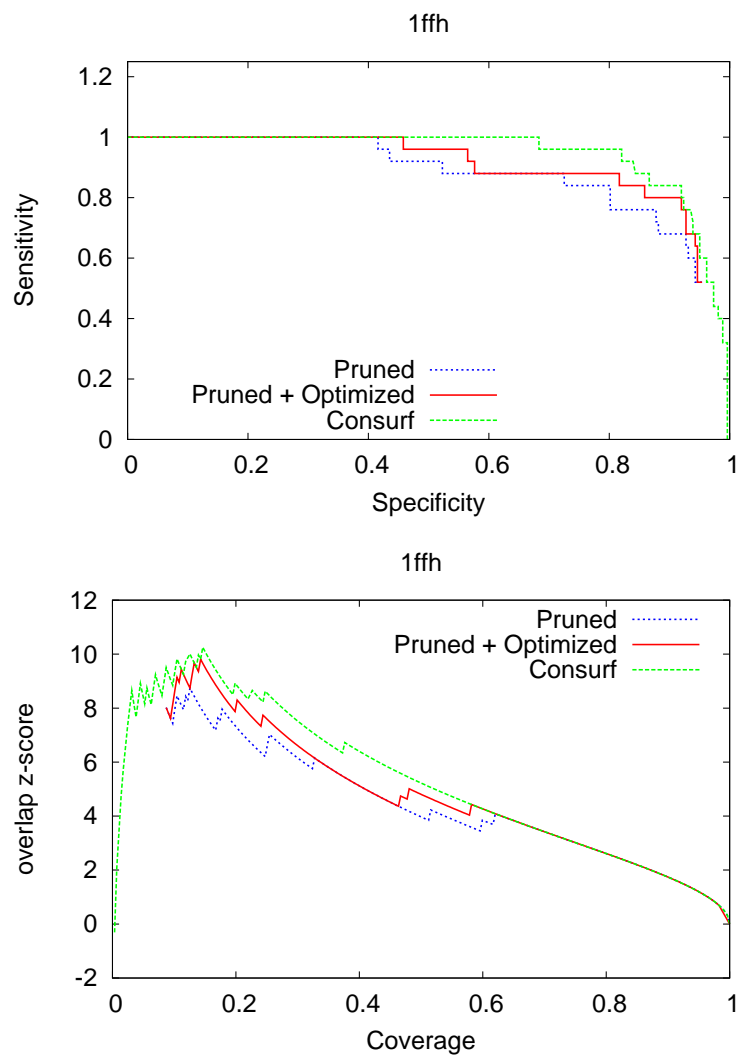
Figure 103: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1emn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
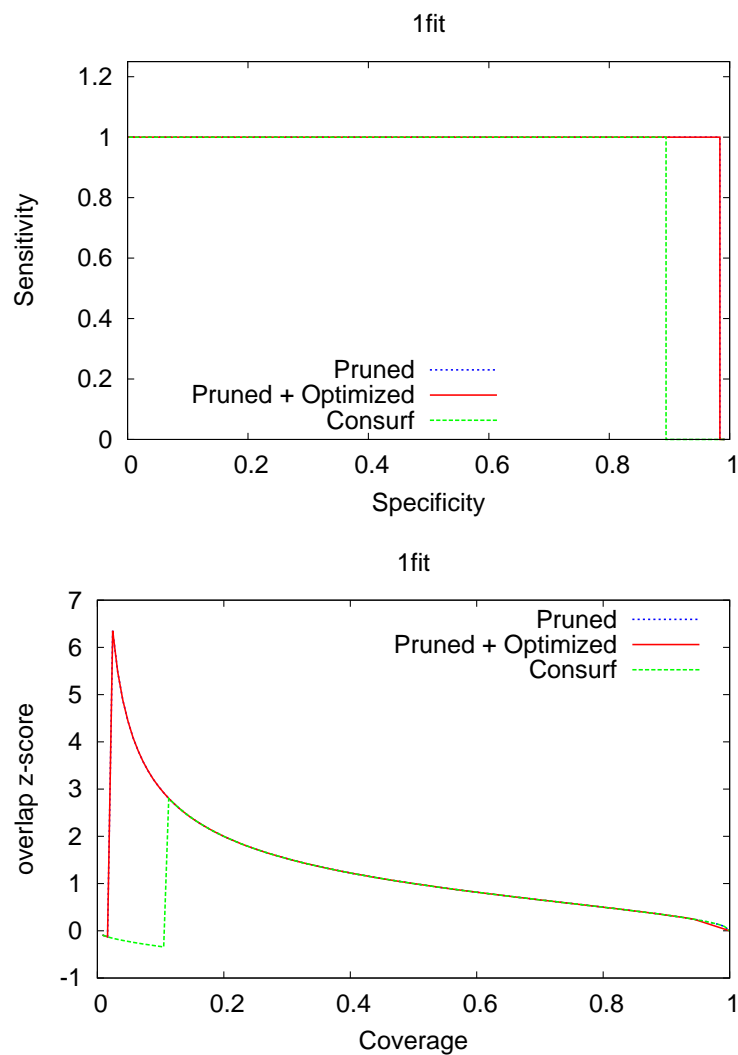
Figure 104: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1esl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
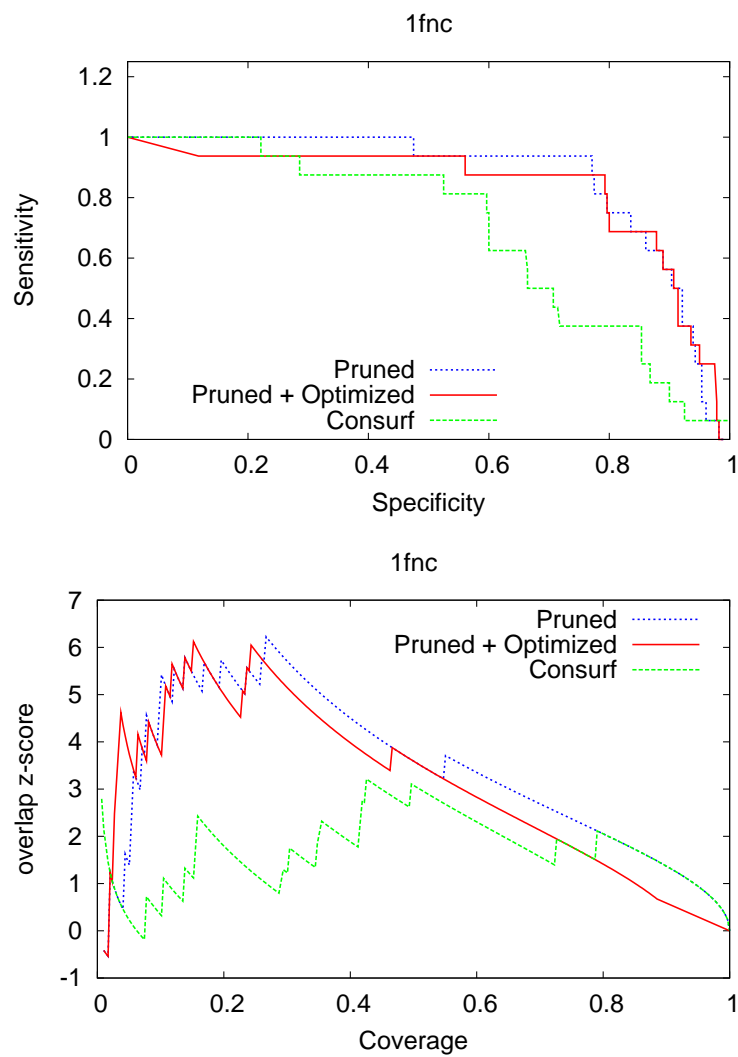
Figure 105: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1far. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
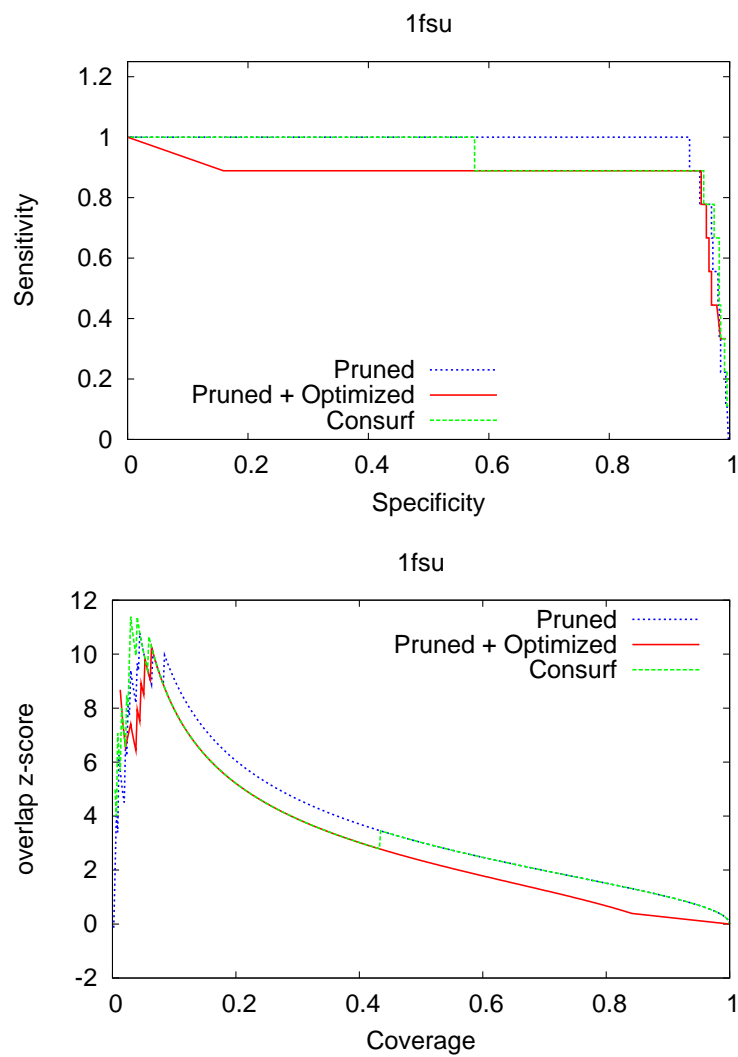
Figure 106: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fca. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
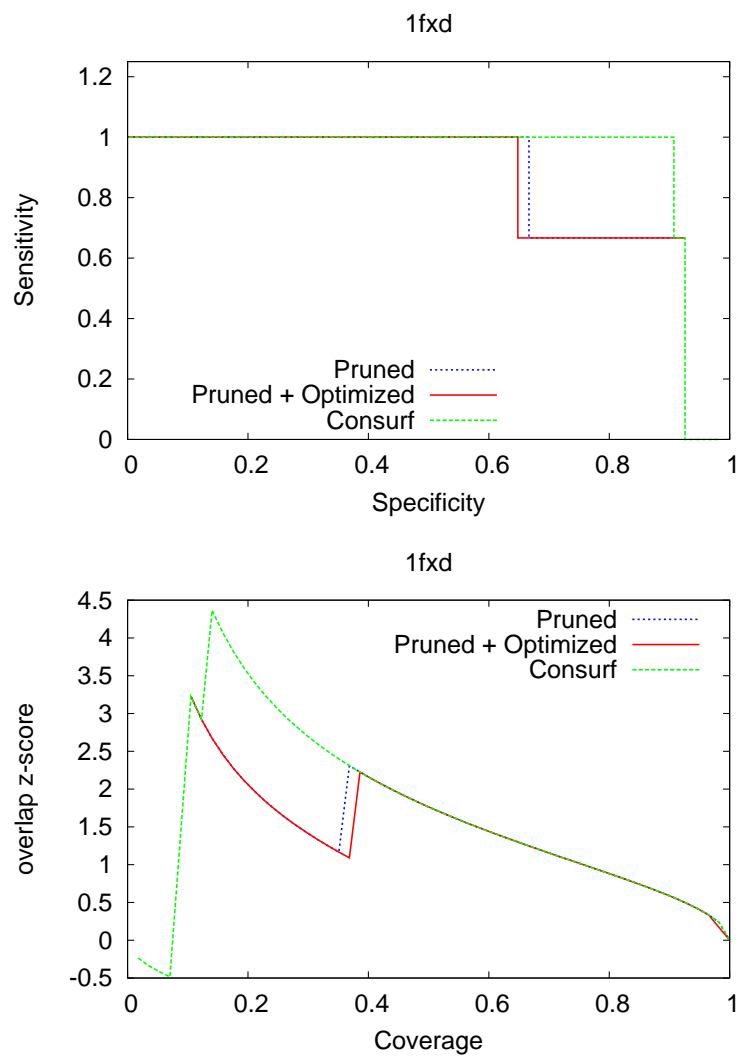
Figure 107: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ffh. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
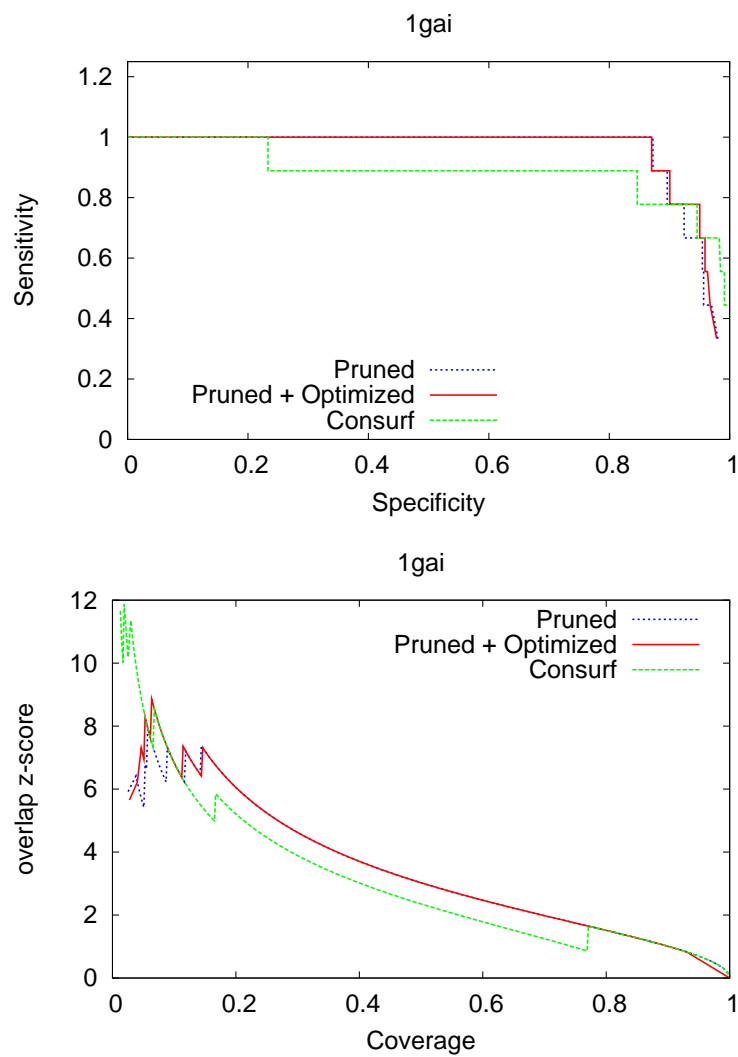
Figure 108: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fit. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
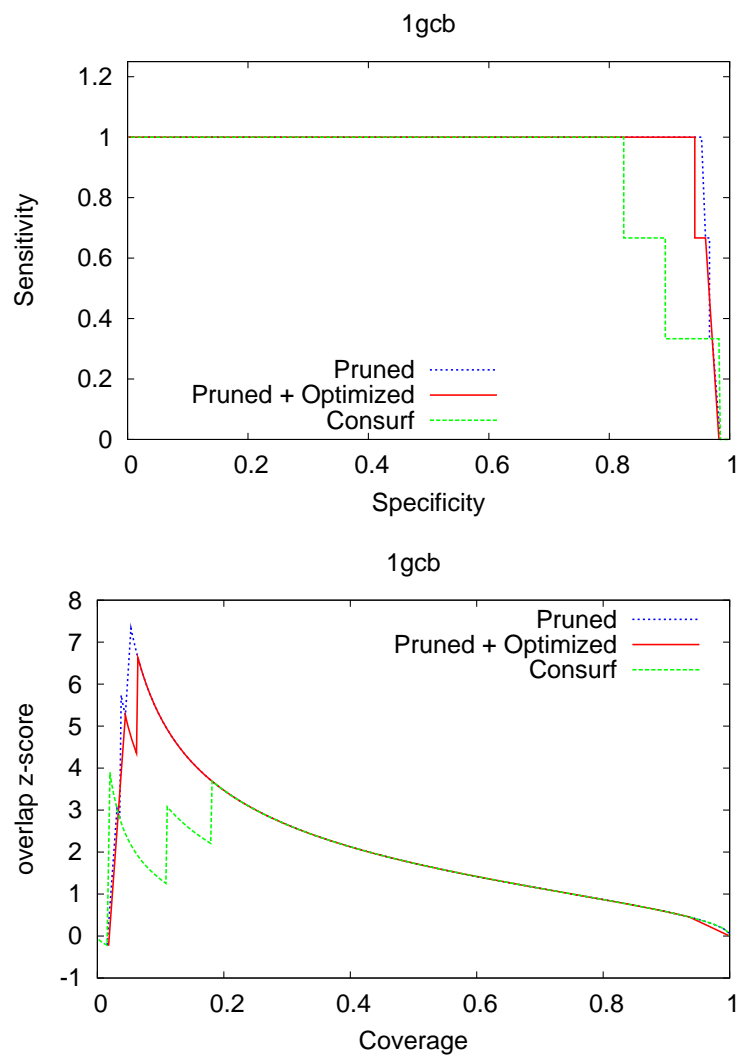
Figure 109: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fnc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
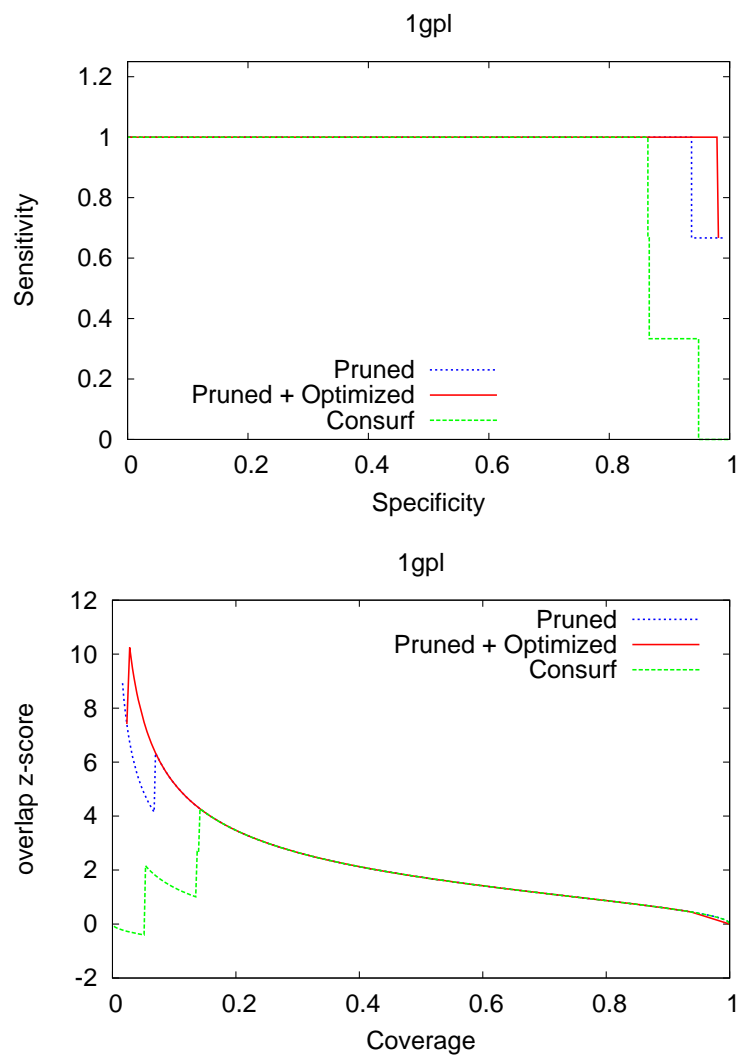
Figure 110: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fsu. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
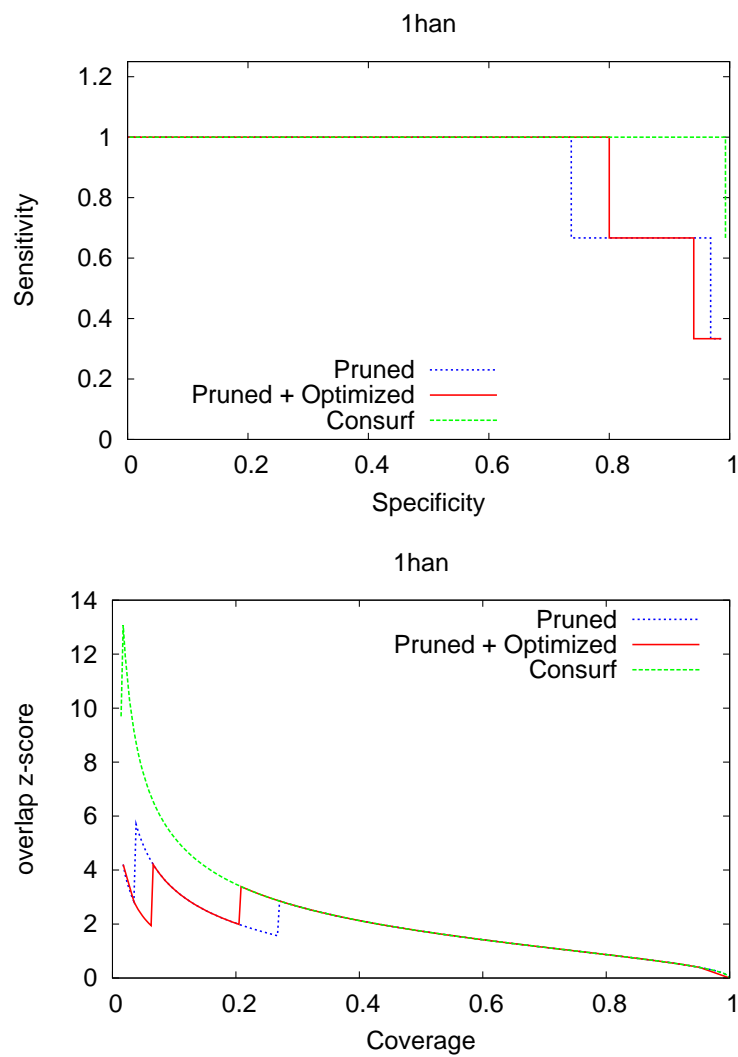
Figure 111: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1fxd. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
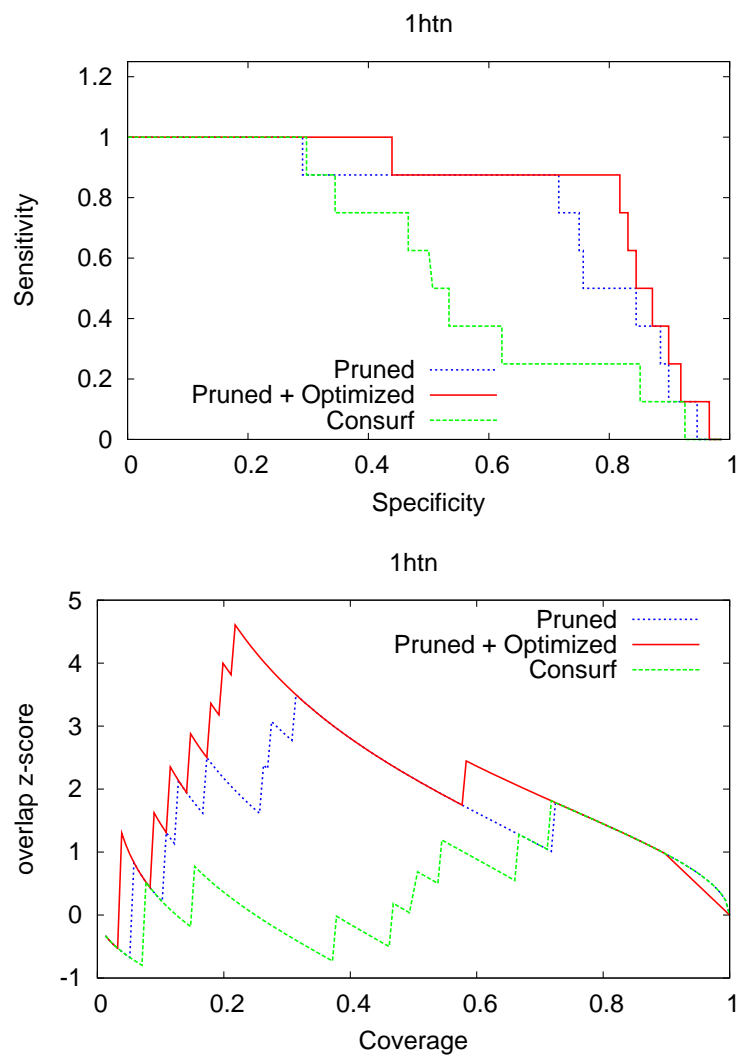
Figure 112: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1gai. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 113: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1gcb. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
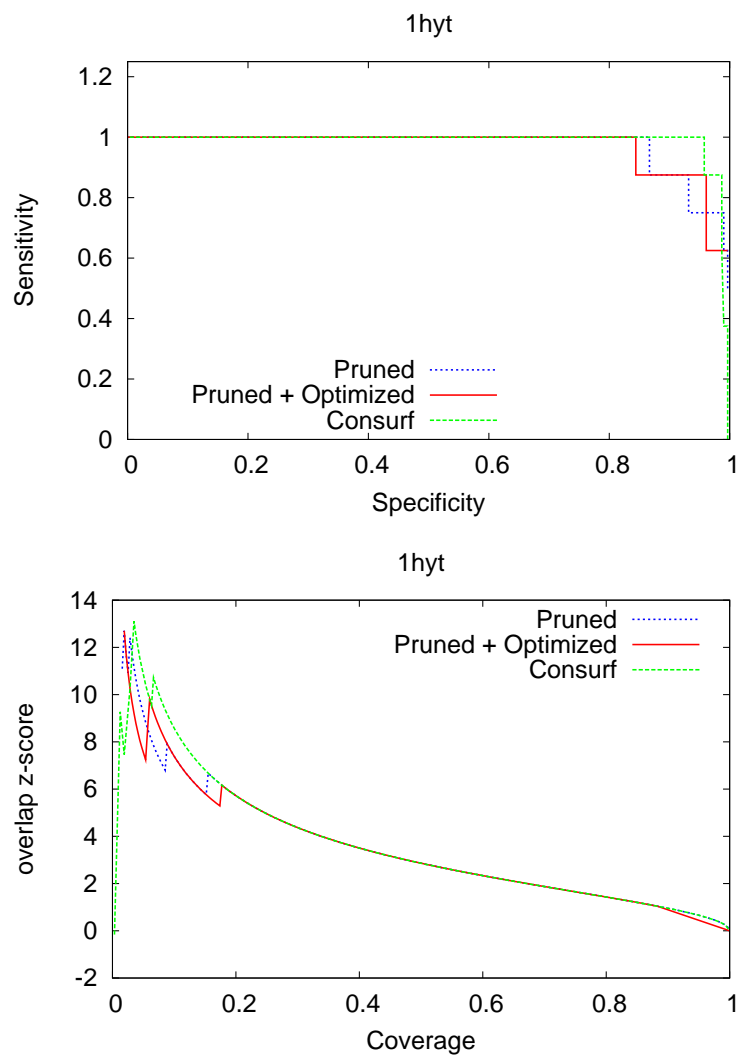
Figure 114: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1gpl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
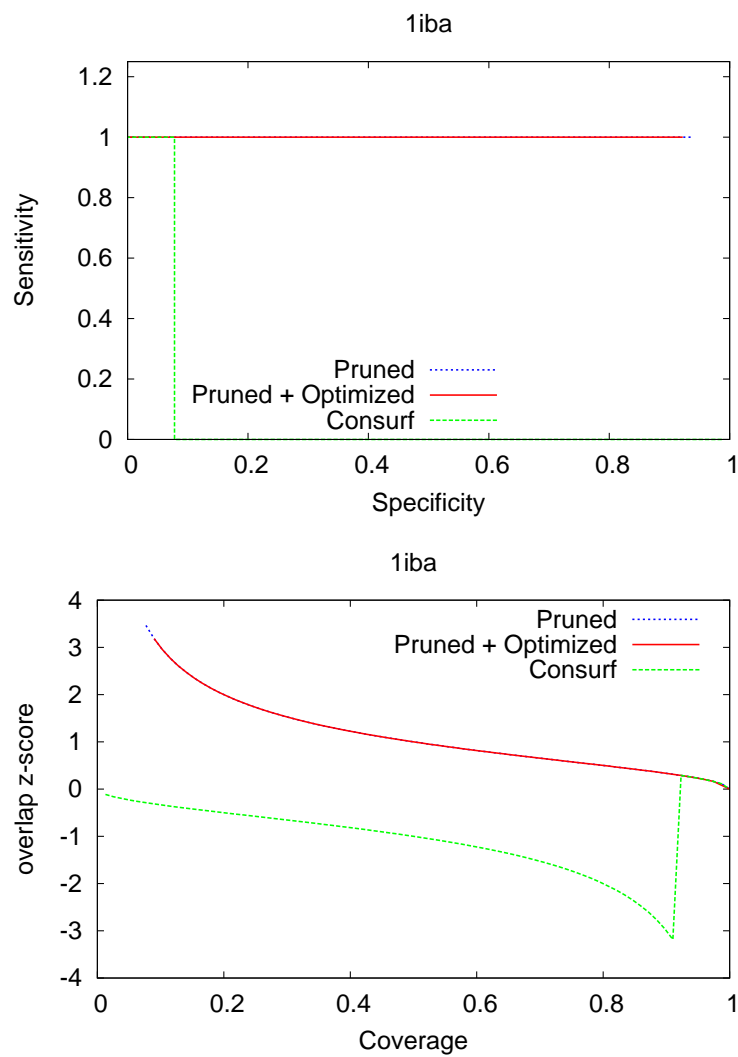
Figure 115: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1han. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
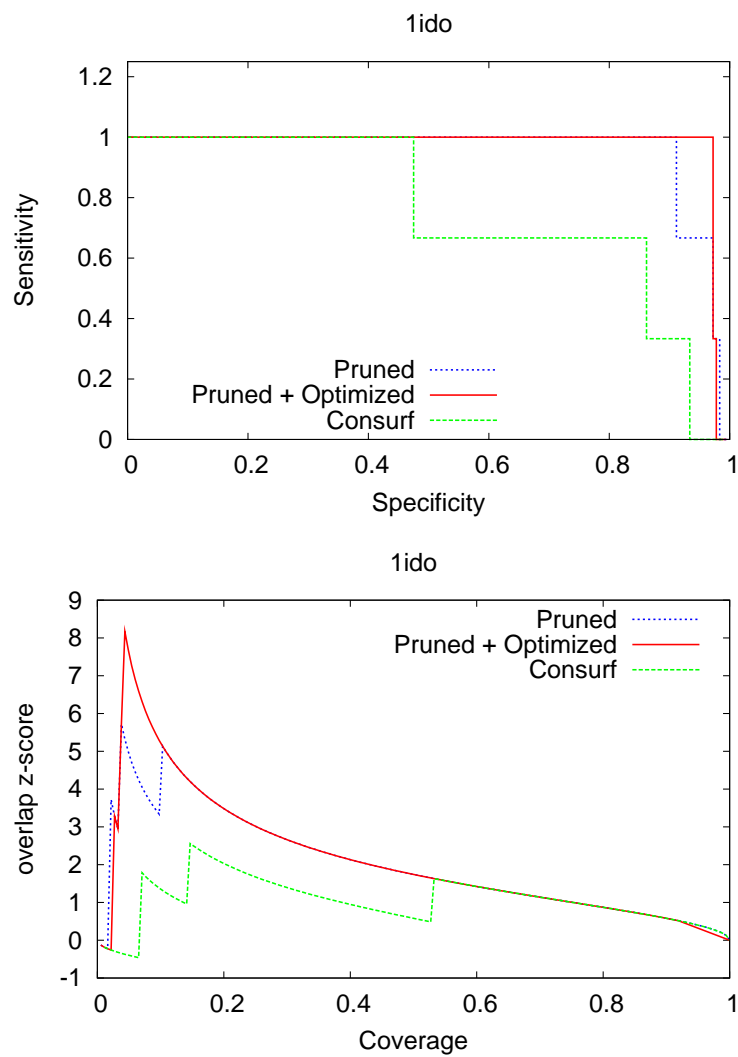
Figure 116: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1htn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
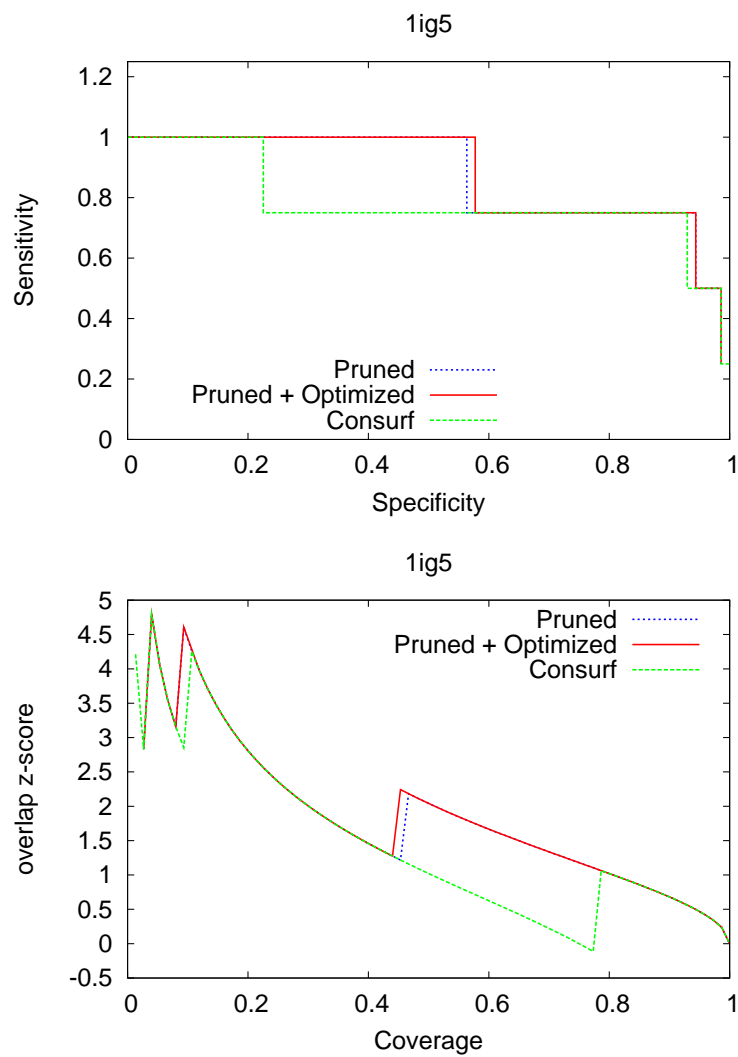
Figure 117: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1hyt. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
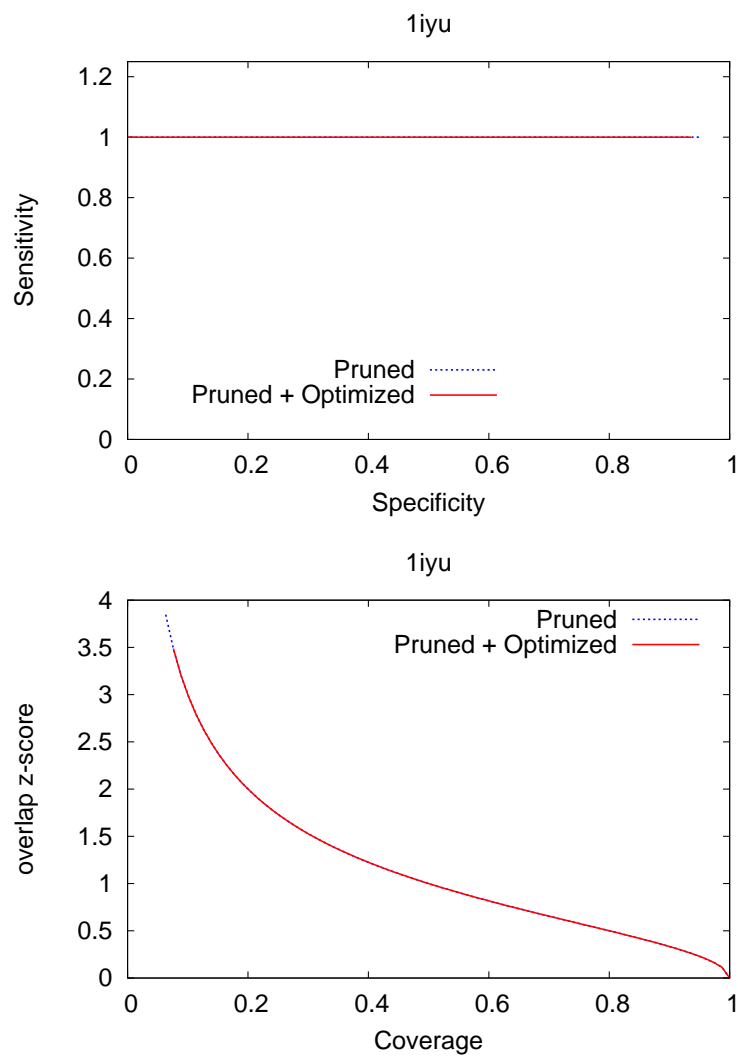
Figure 118: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1iba. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
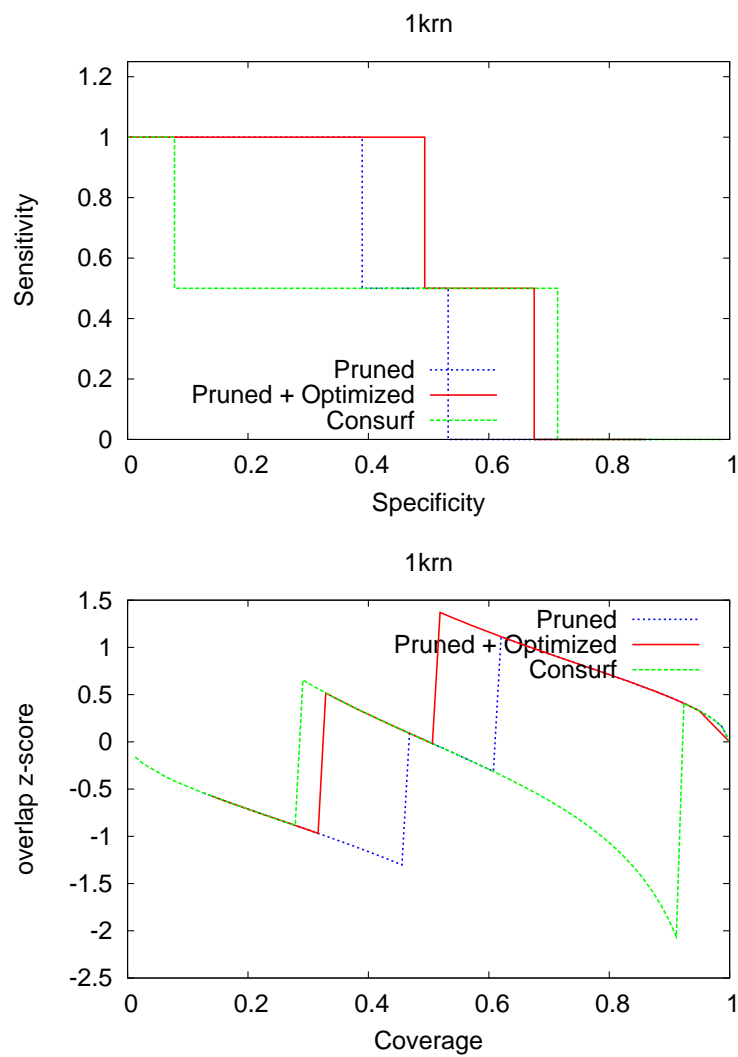
Figure 119: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ido. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
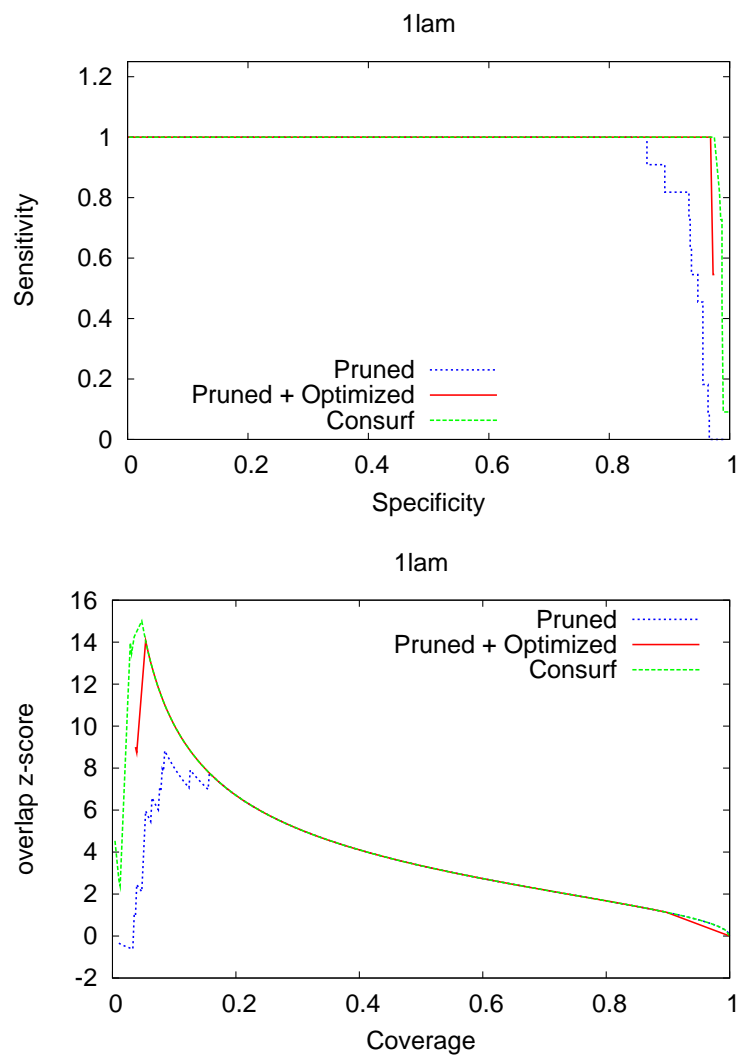
Figure 120: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1ig5. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
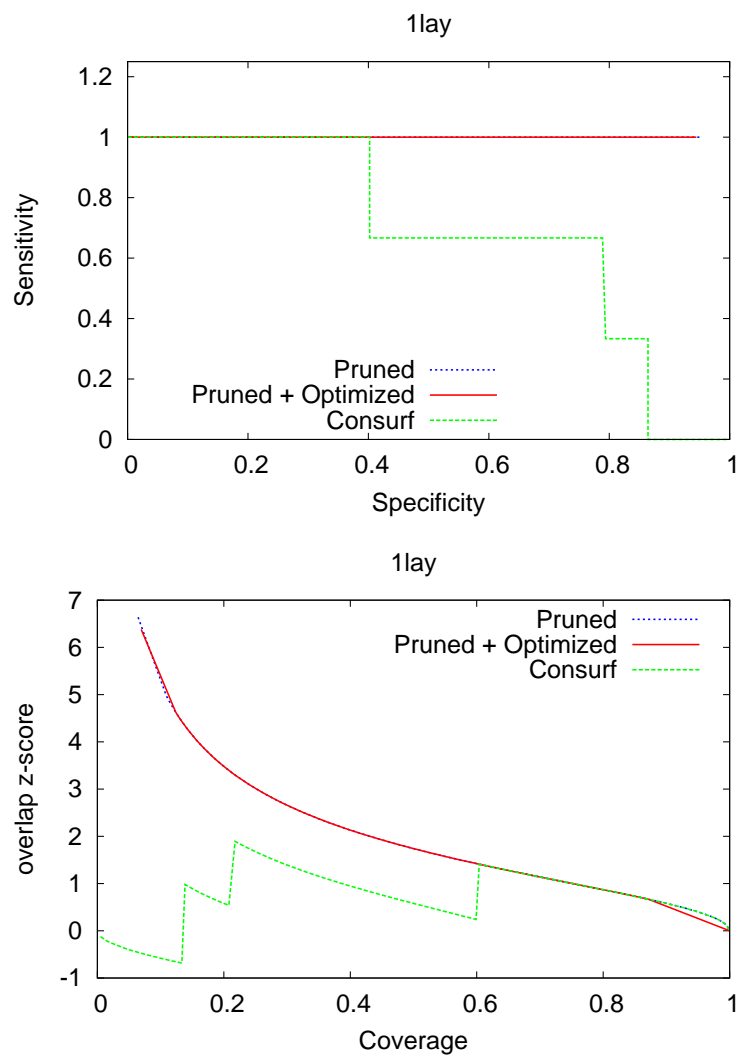
Figure 121: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1iyu. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
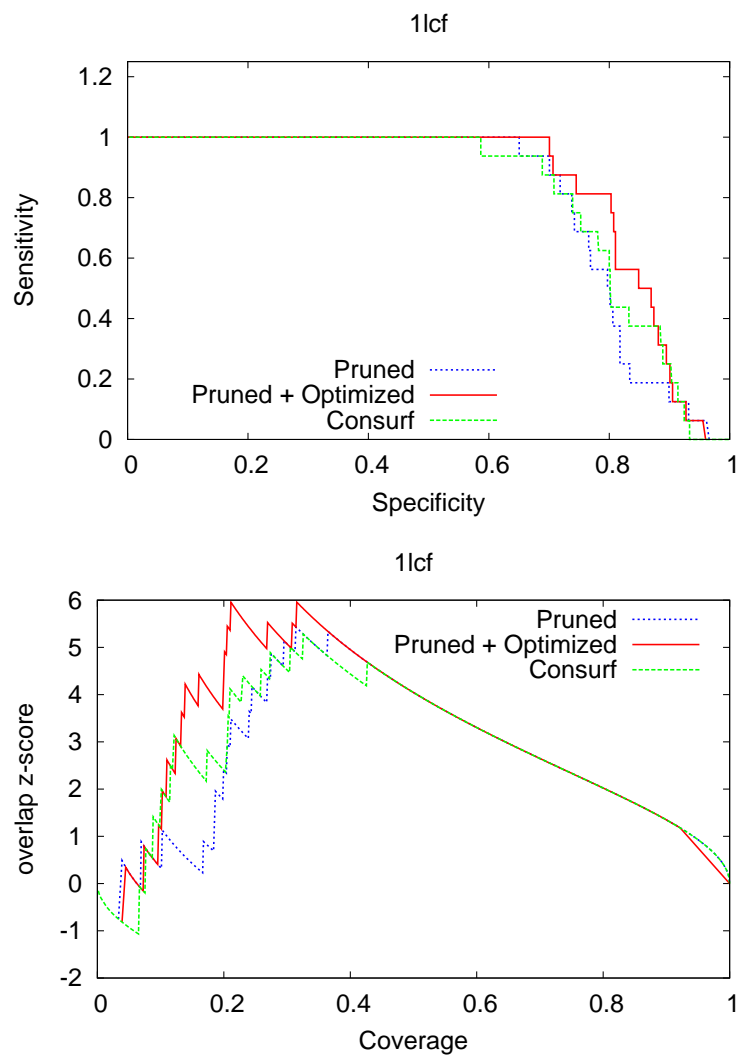
Figure 122: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1krn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
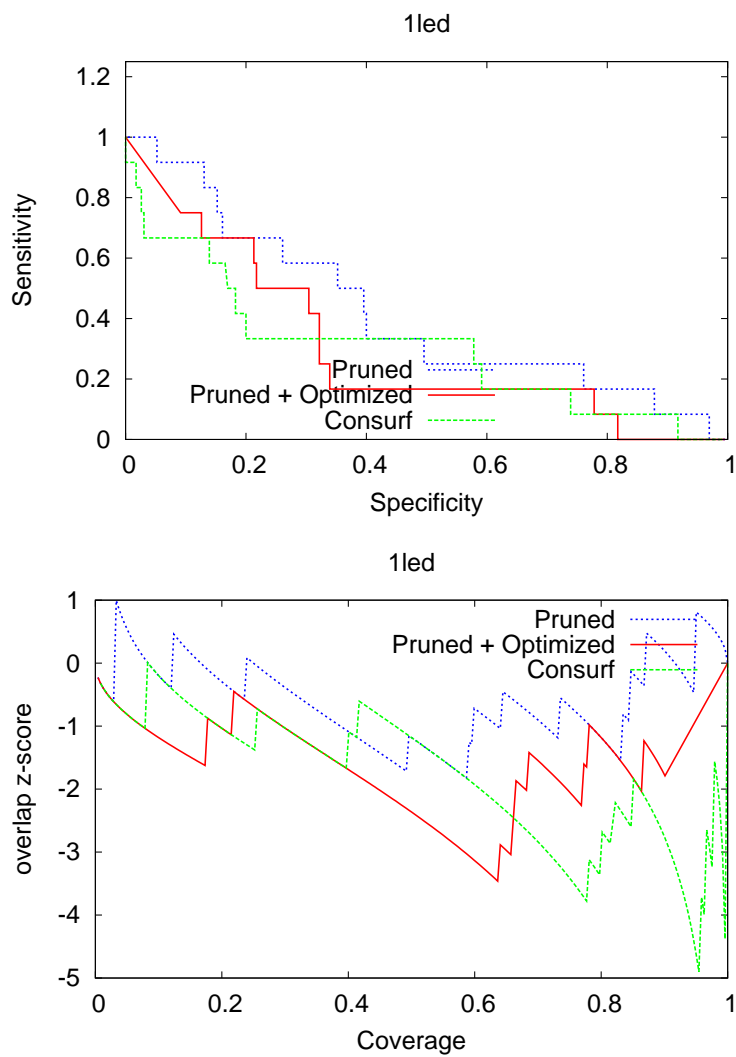
Figure 123: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1lam. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
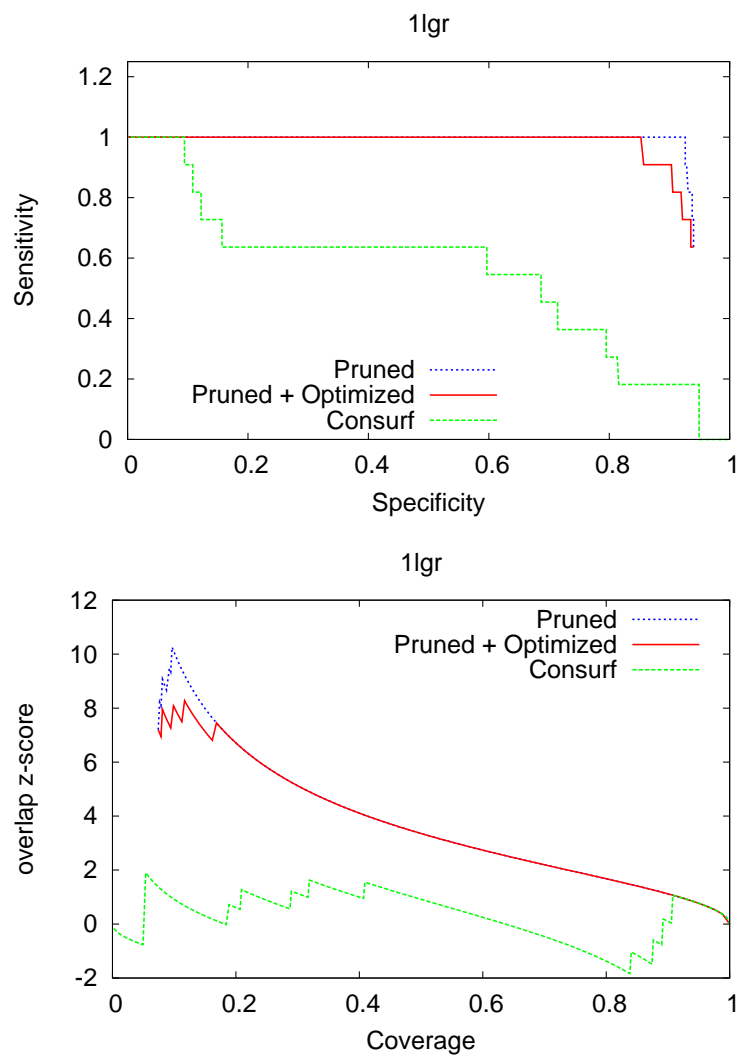
Figure 124: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1lay. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
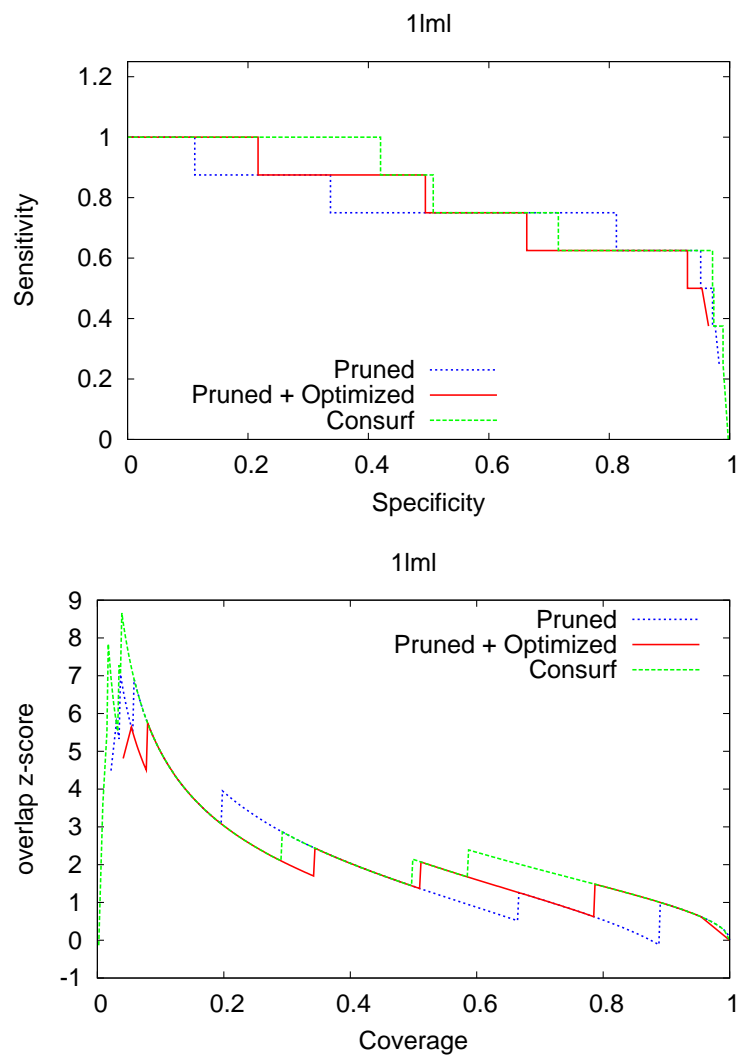
Figure 125: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1lcf. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
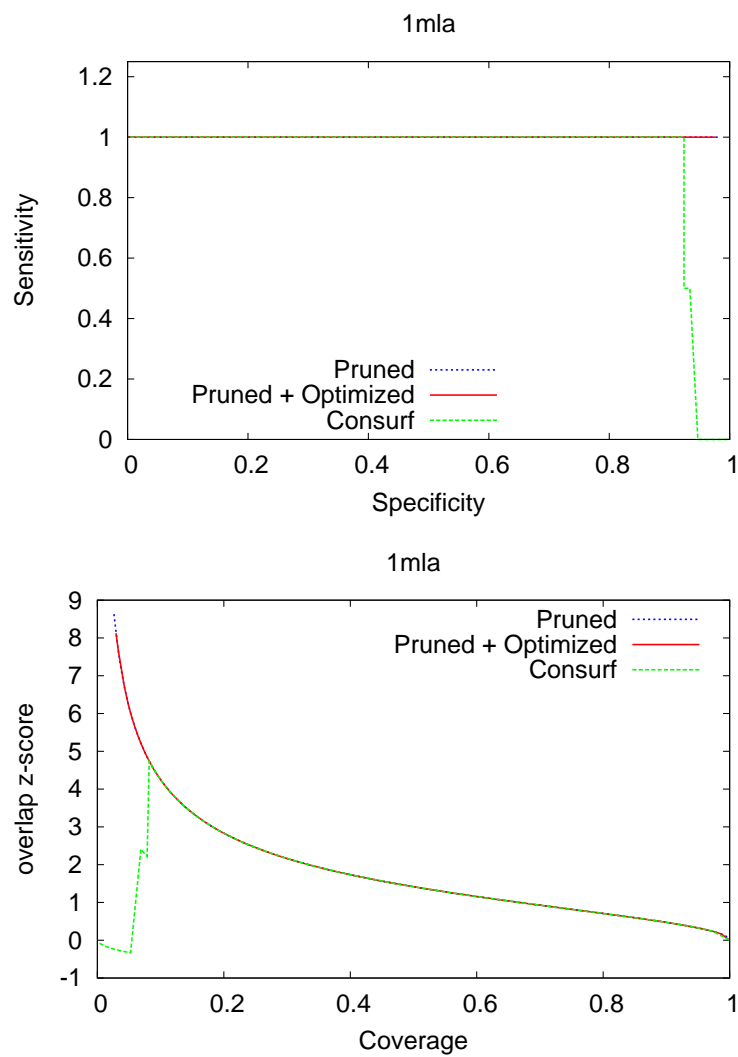
Figure 126: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1led. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
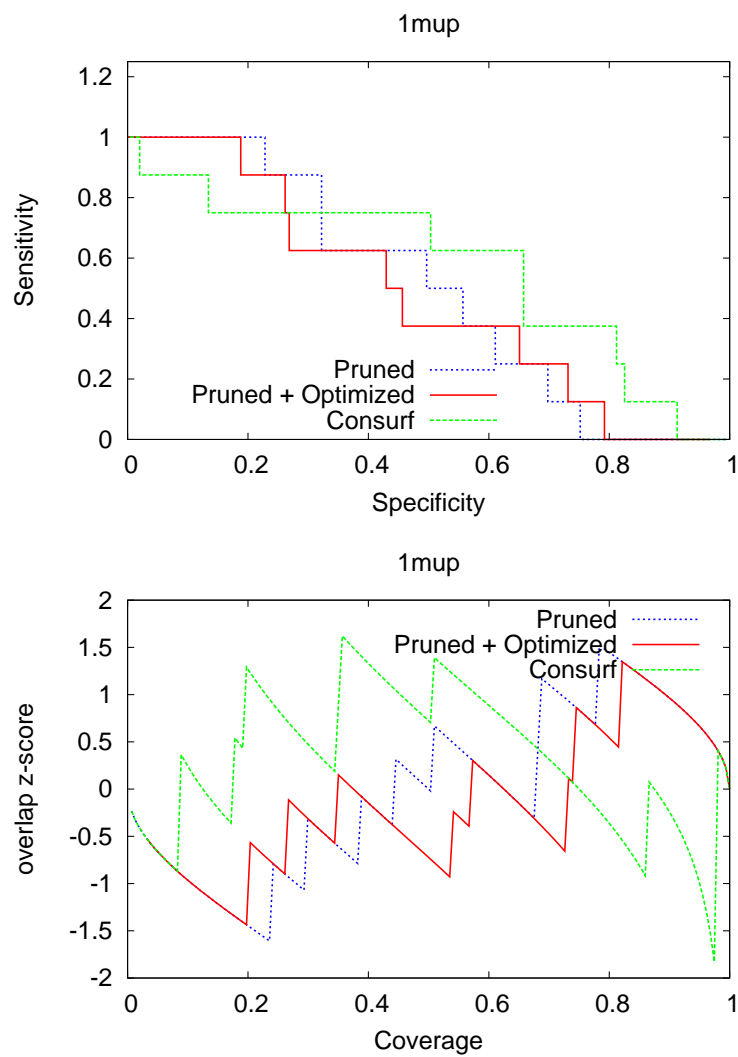
Figure 127: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1lgr. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
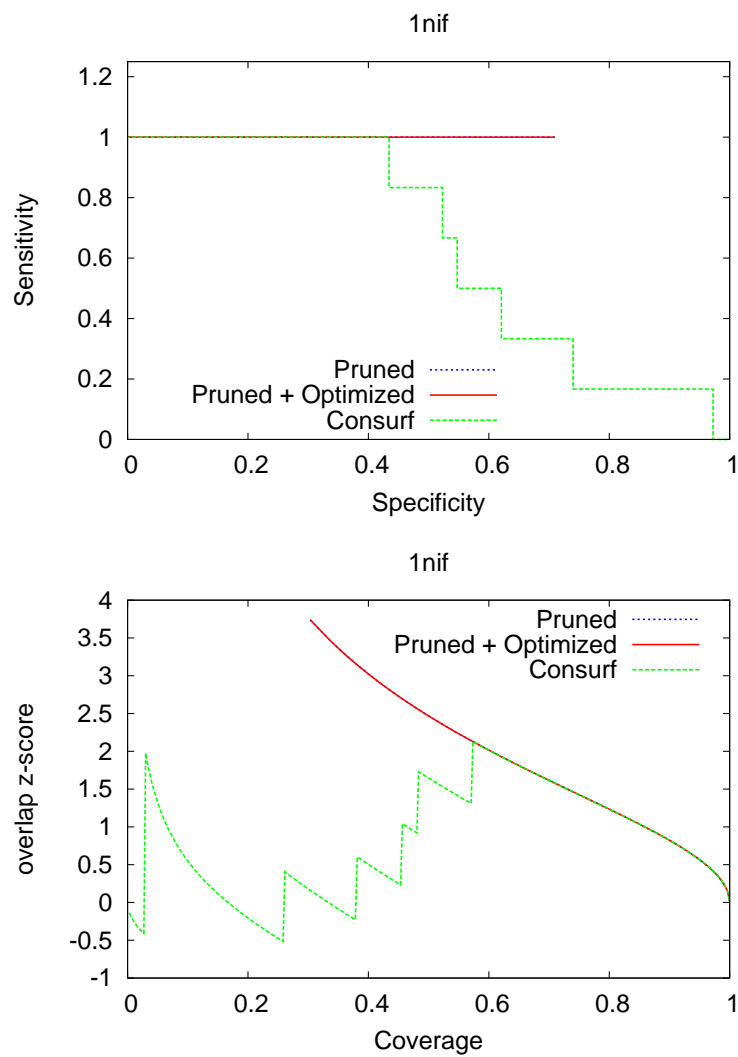
Figure 128: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1lml. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
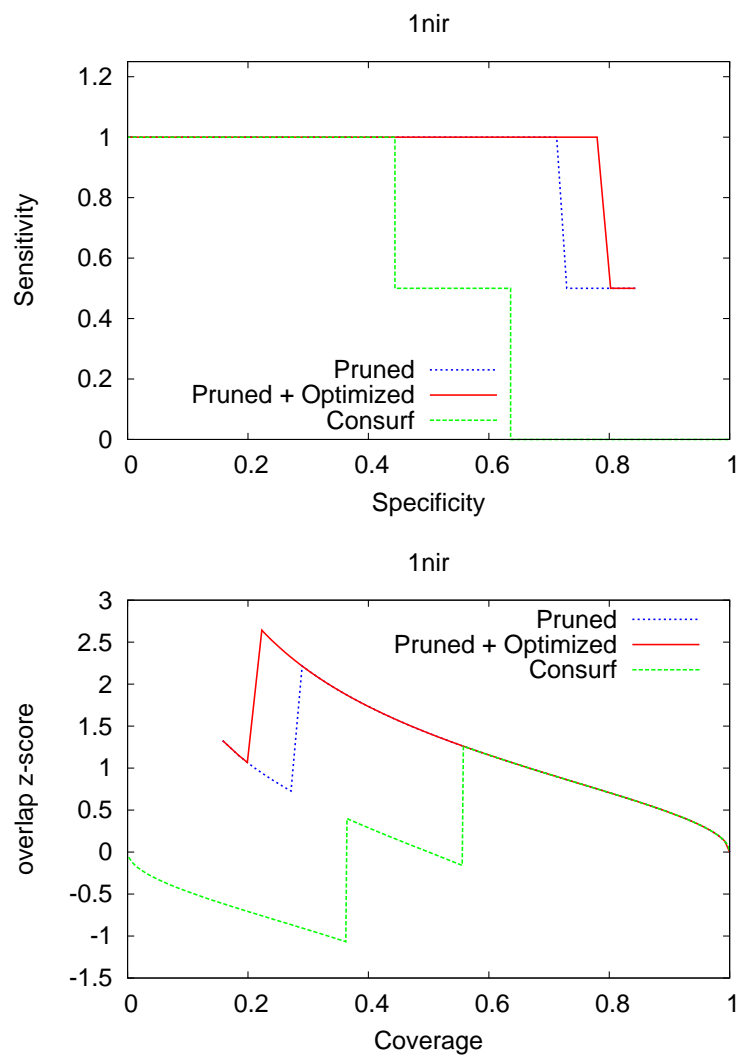
Figure 129: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1mla. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 130: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1mup. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
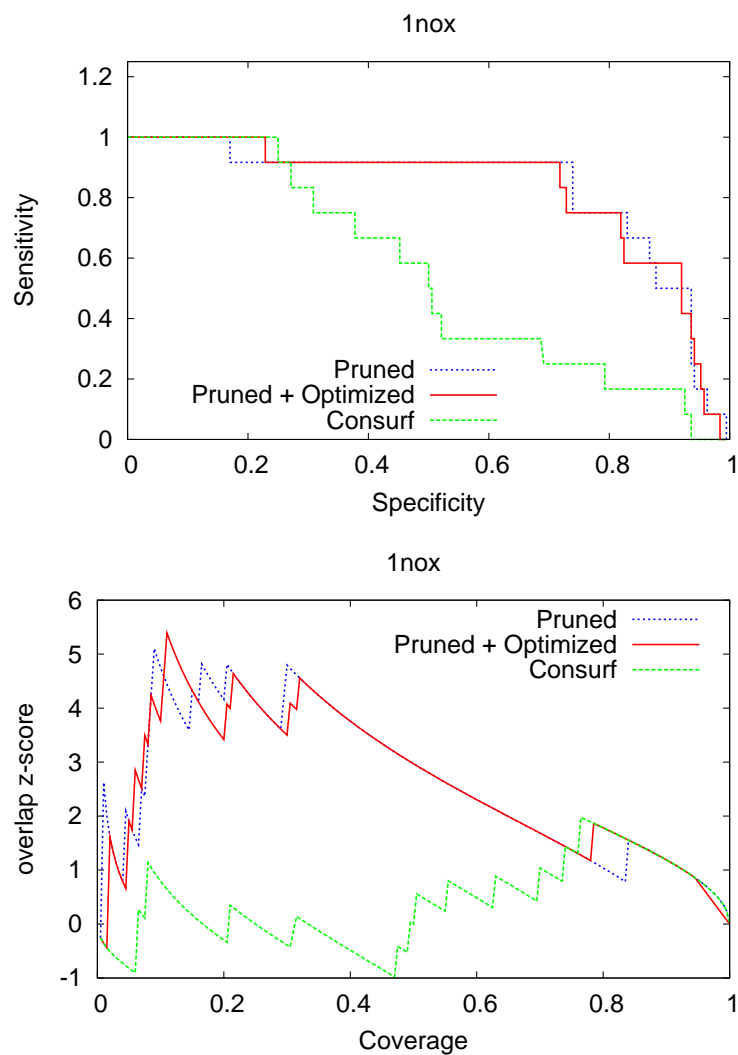
Figure 131: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1nif. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
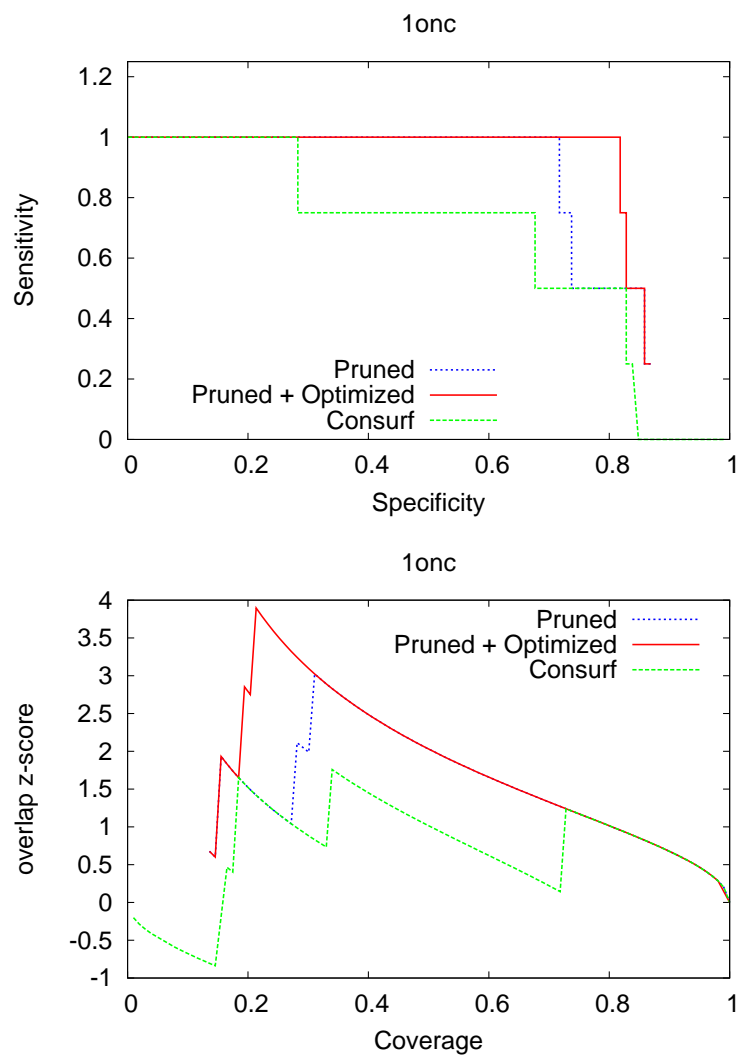
Figure 132: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1nir. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
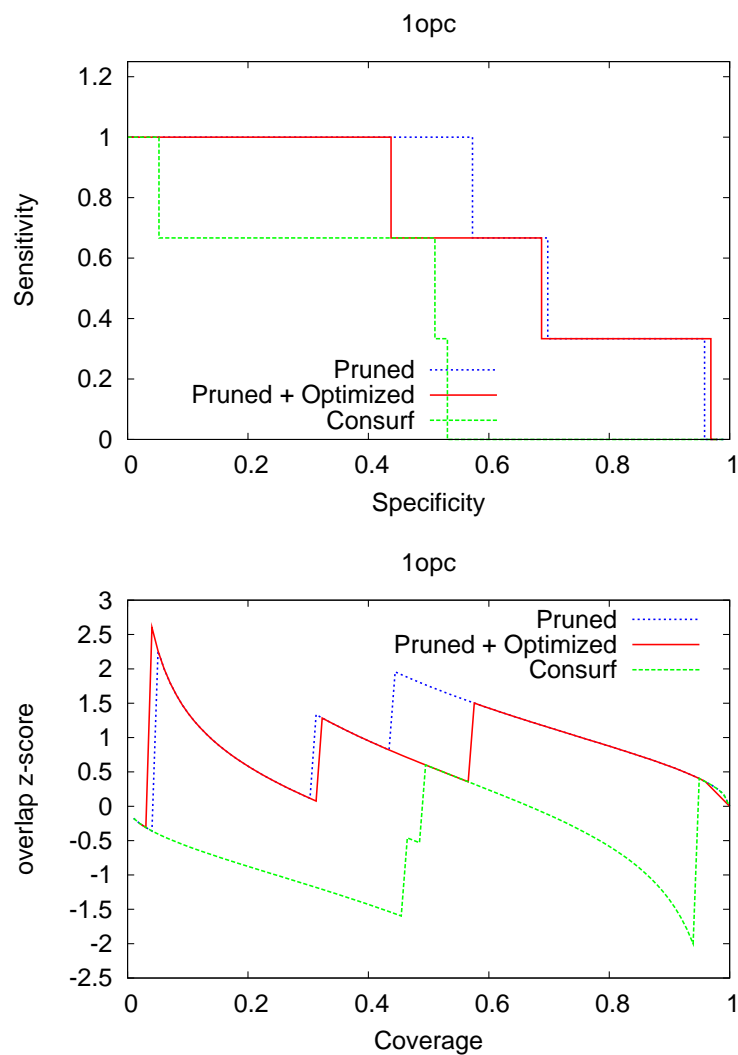
Figure 133: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1nox. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
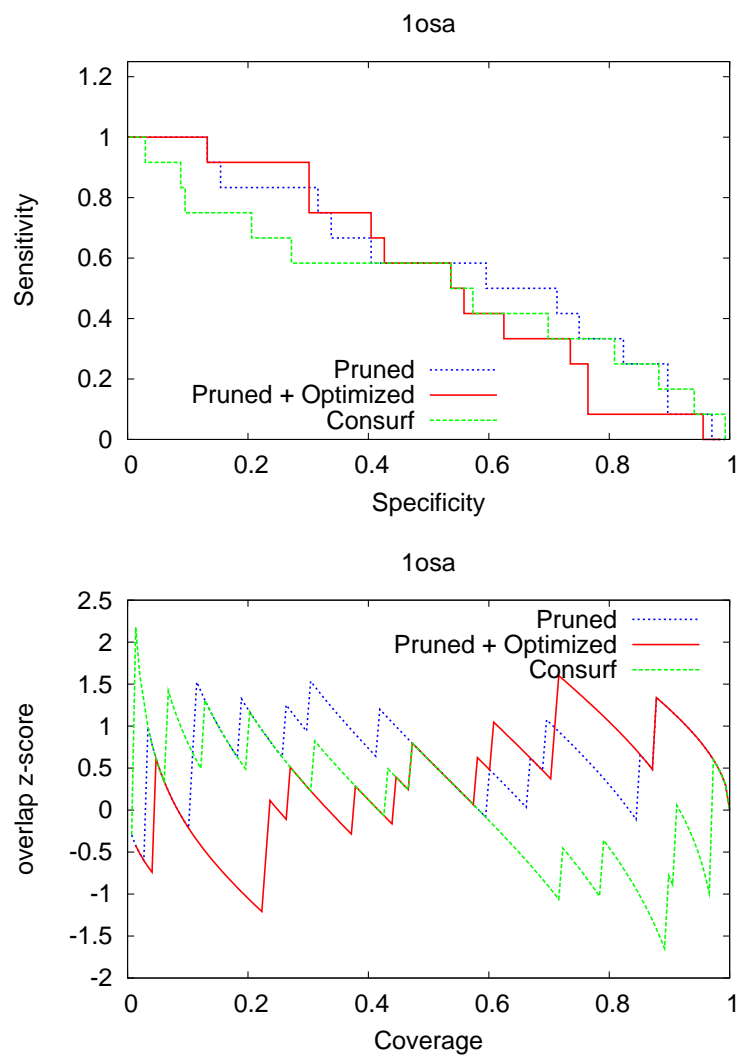
Figure 134: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1onc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
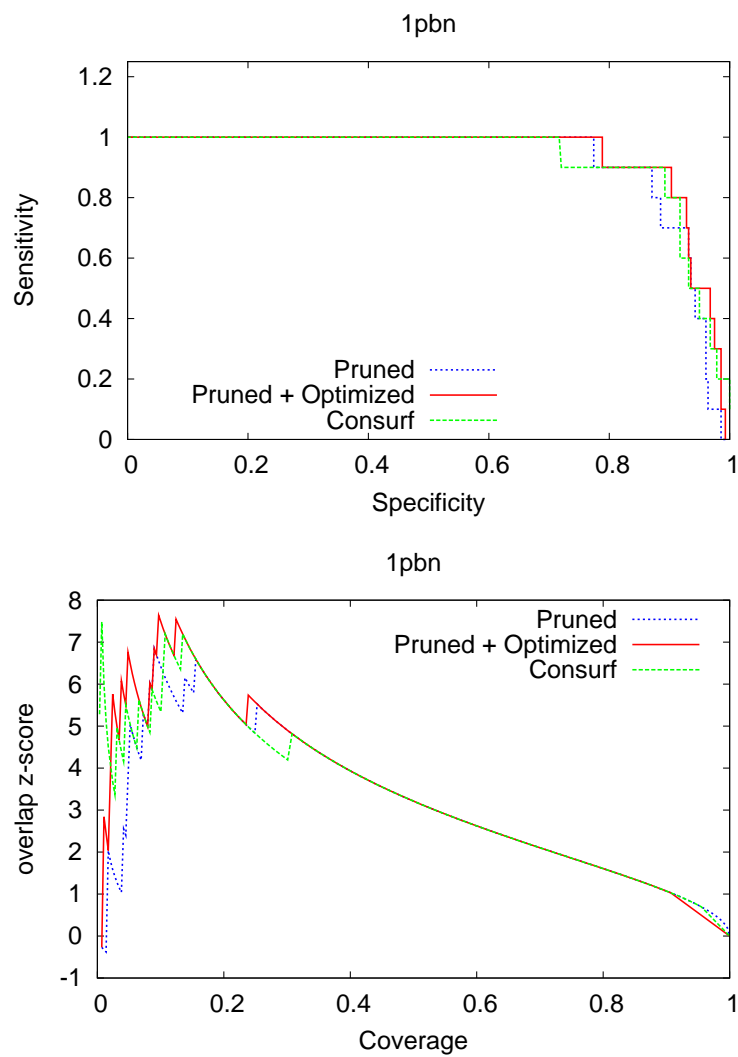
Figure 135: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1opc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
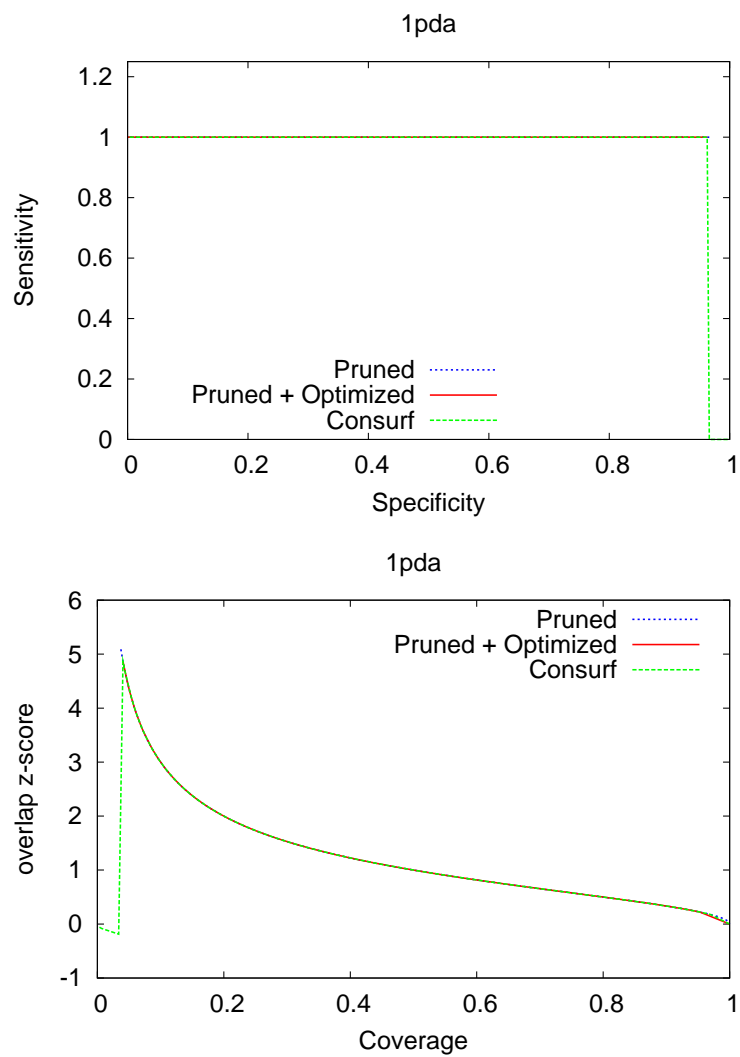
Figure 136: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1osa. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
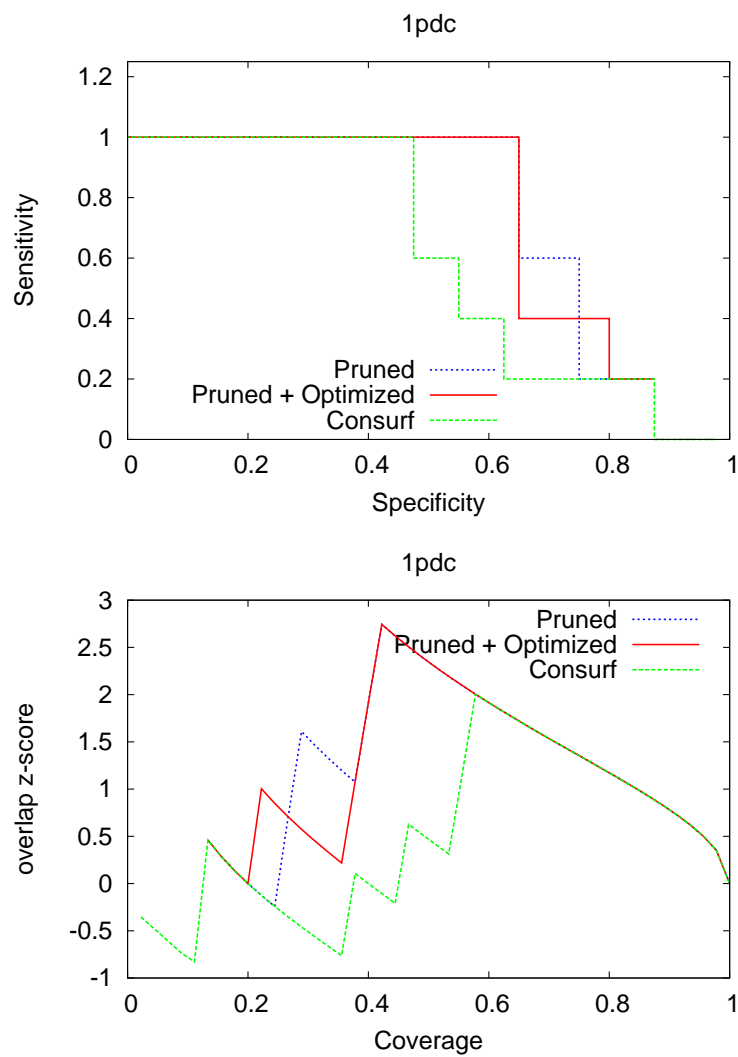
Figure 137: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pbn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
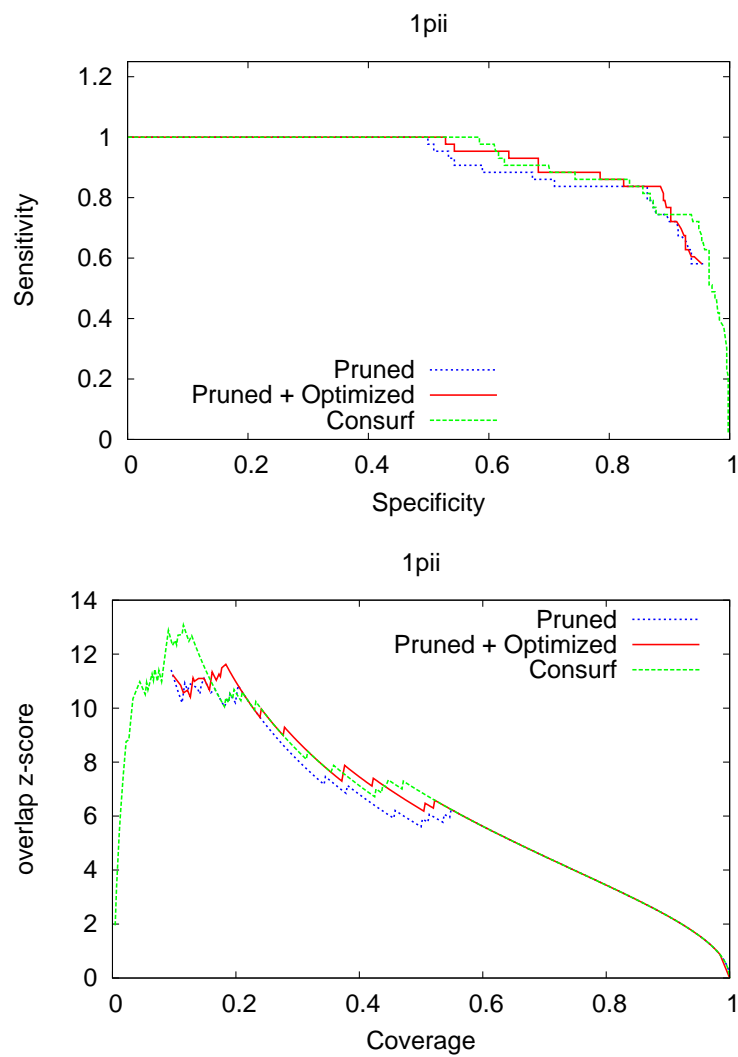
Figure 138: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pda. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
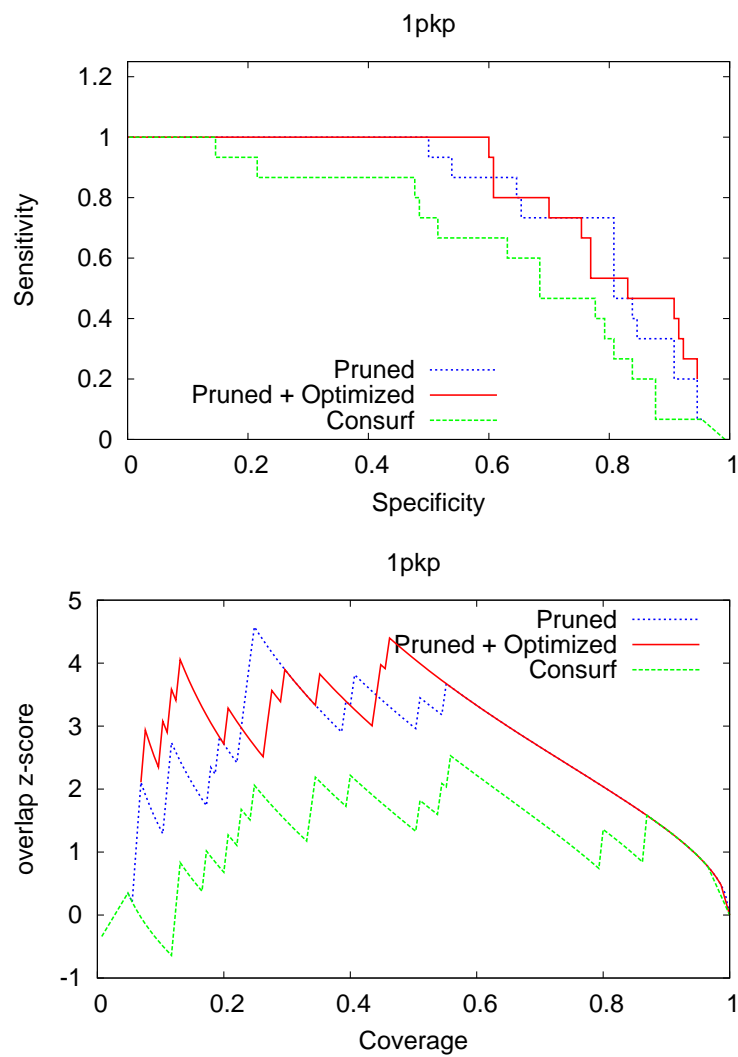
Figure 139: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pdc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
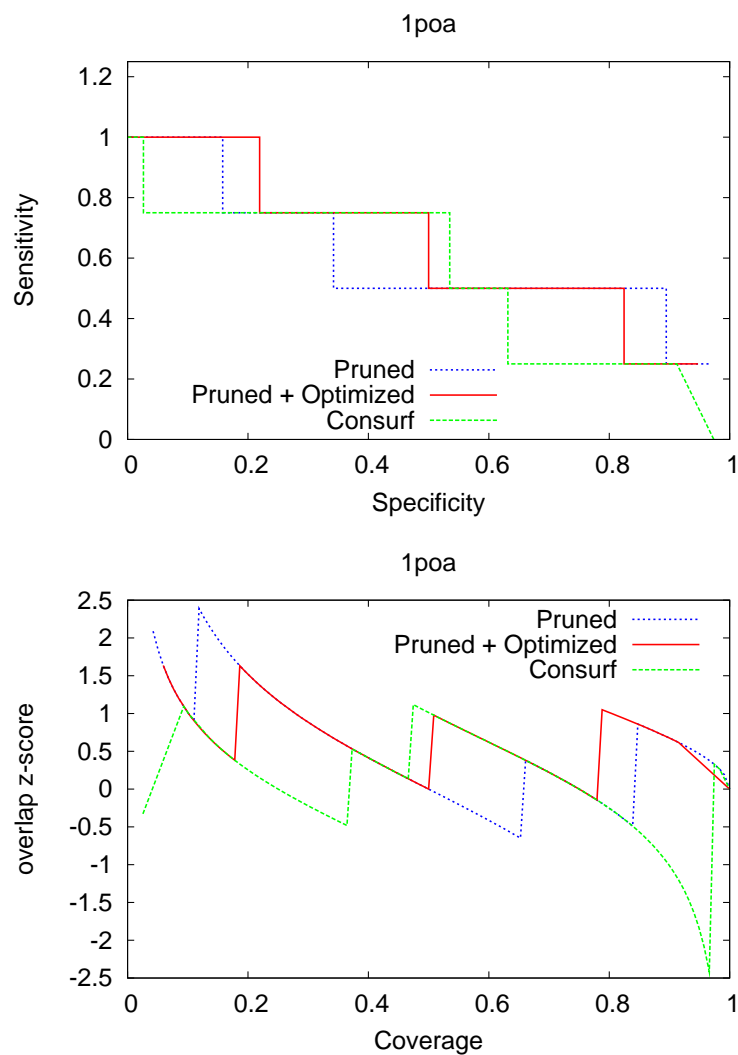
Figure 140: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pii. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
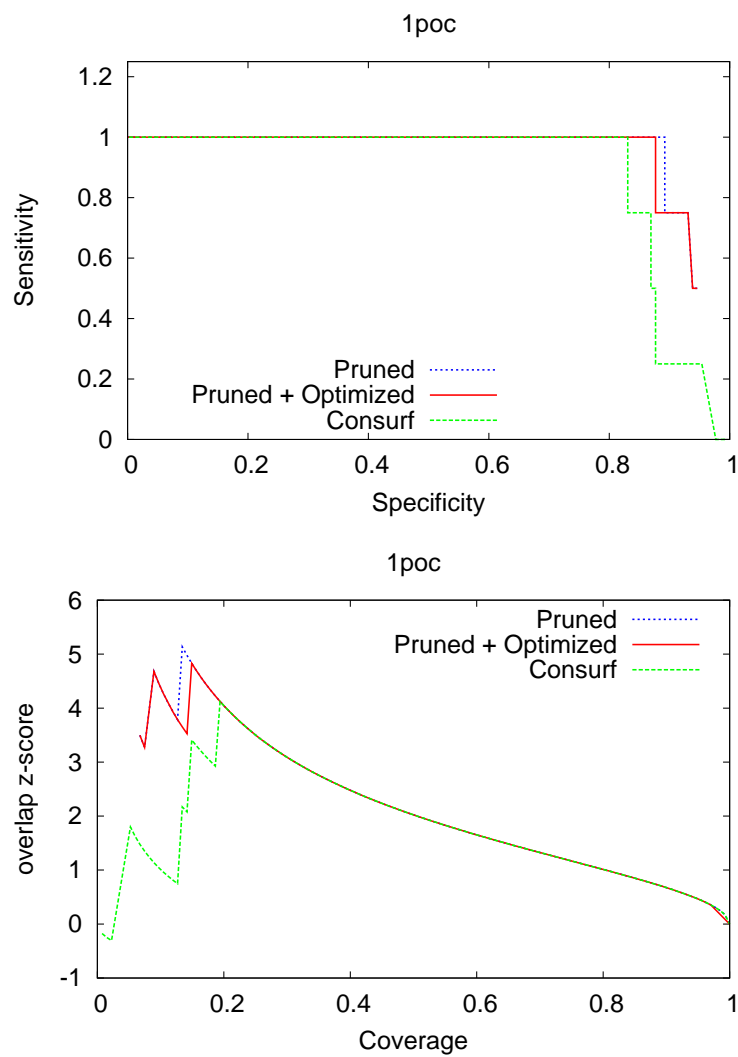
Figure 141: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pkp. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
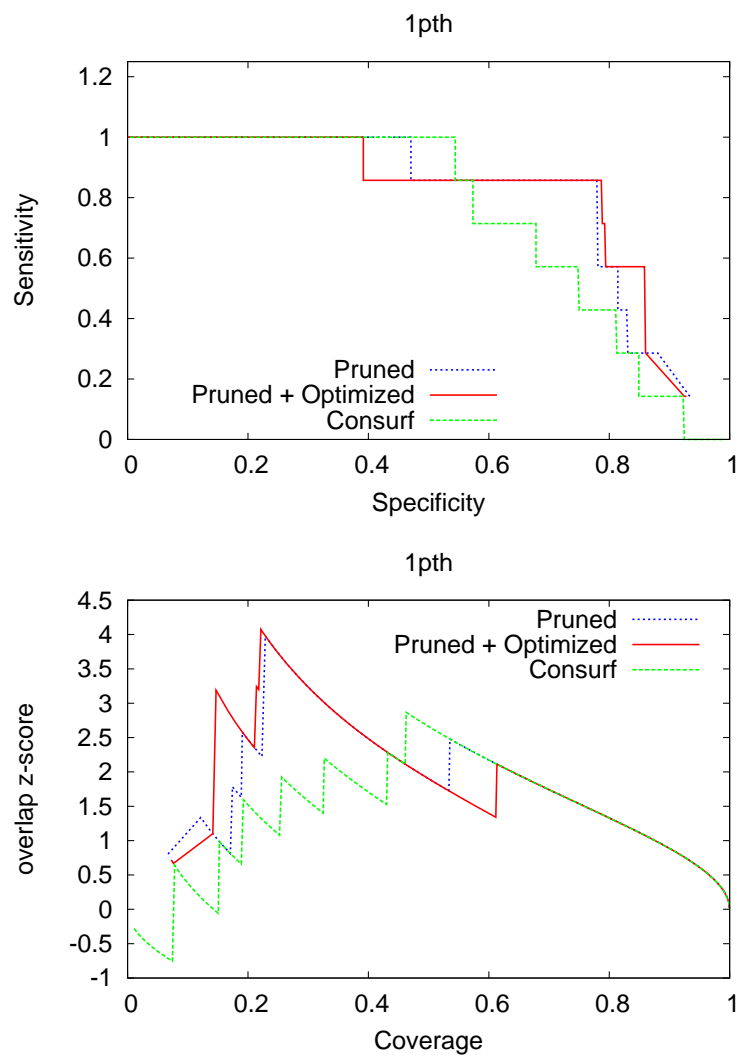
Figure 142: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1poa. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
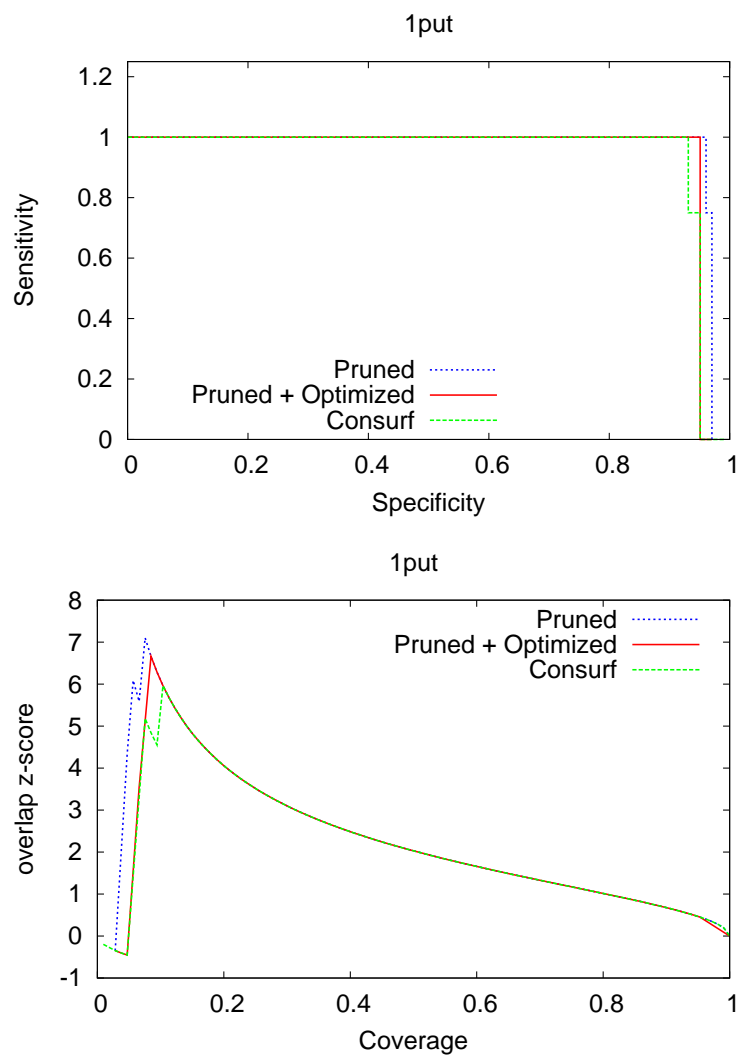
Figure 143: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1poc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
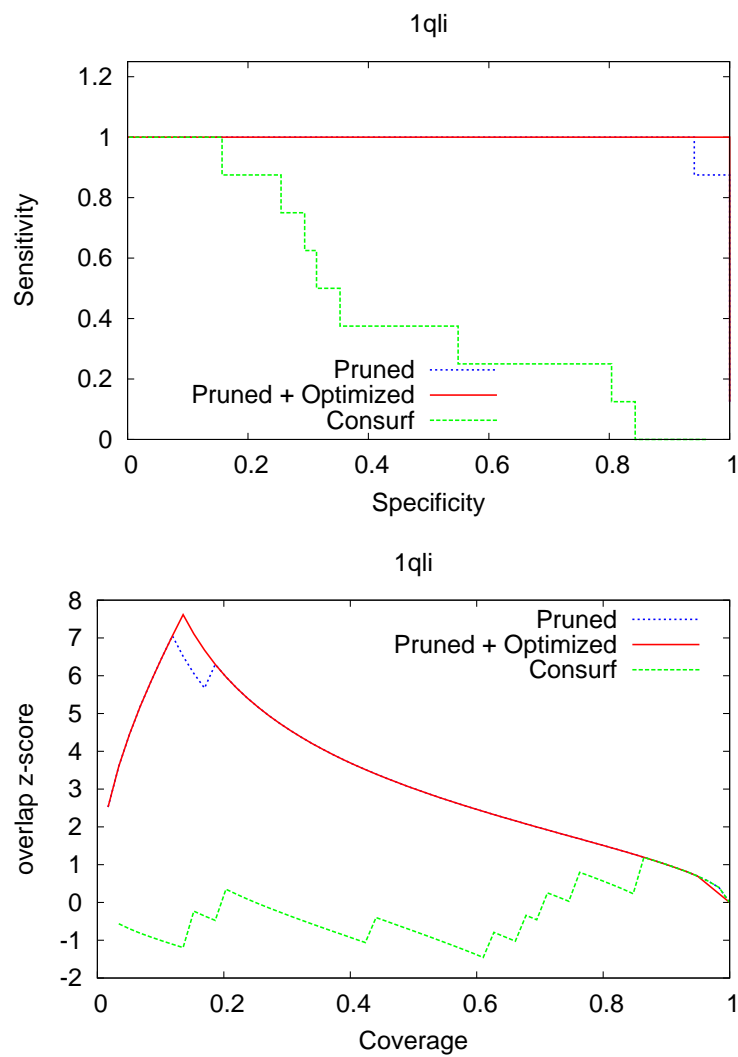
146

Figure 144: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1pth. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
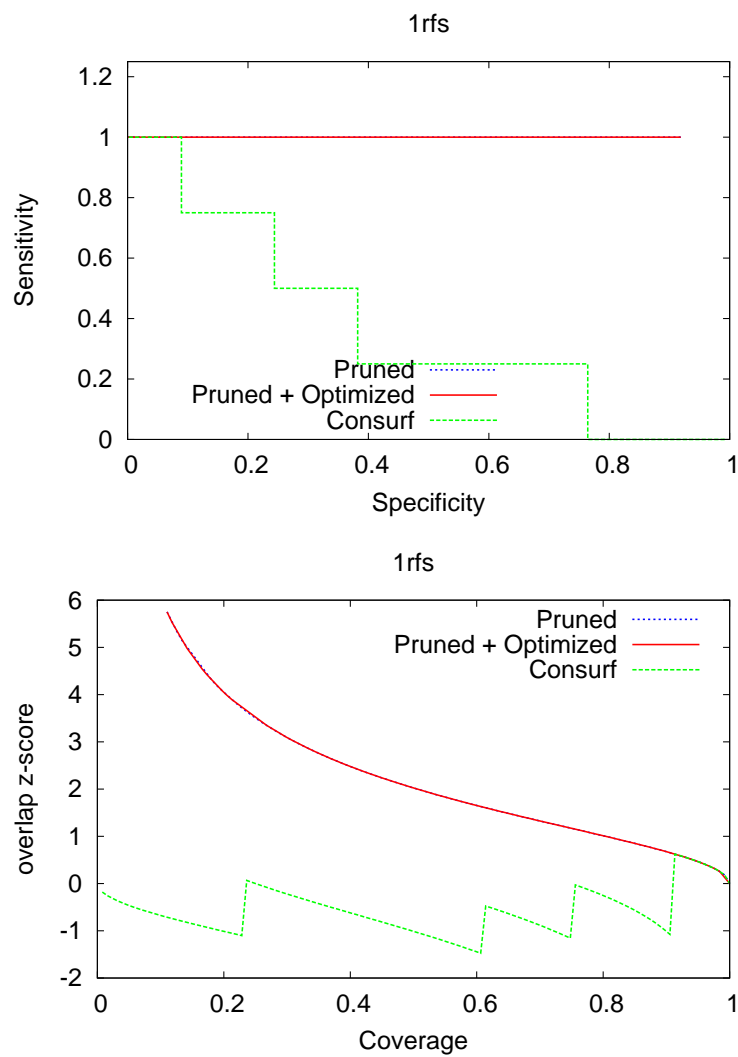
Figure 145: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1put. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
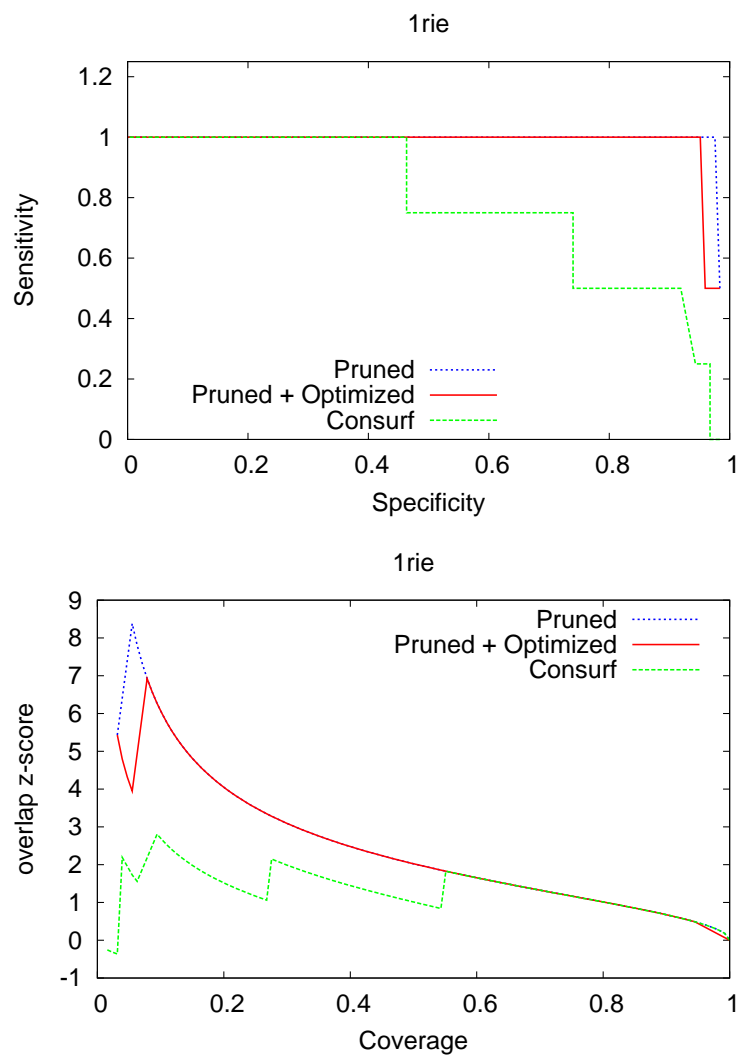
Figure 146: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1qli. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 147: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1rfs. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
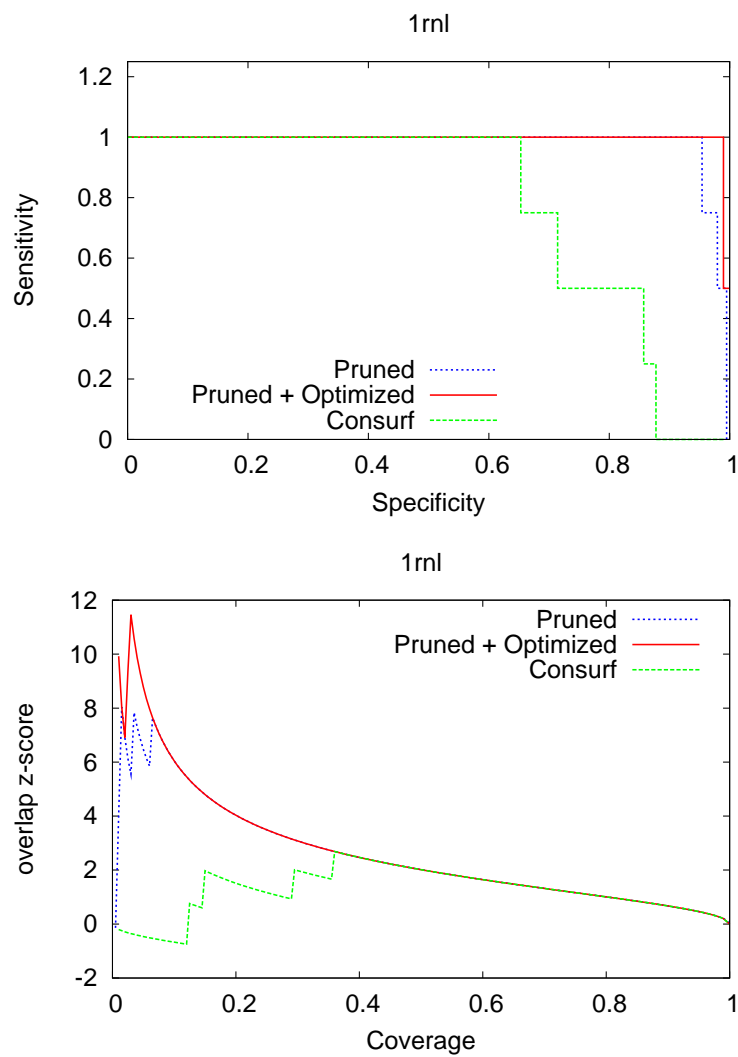
Figure 148: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1rie. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
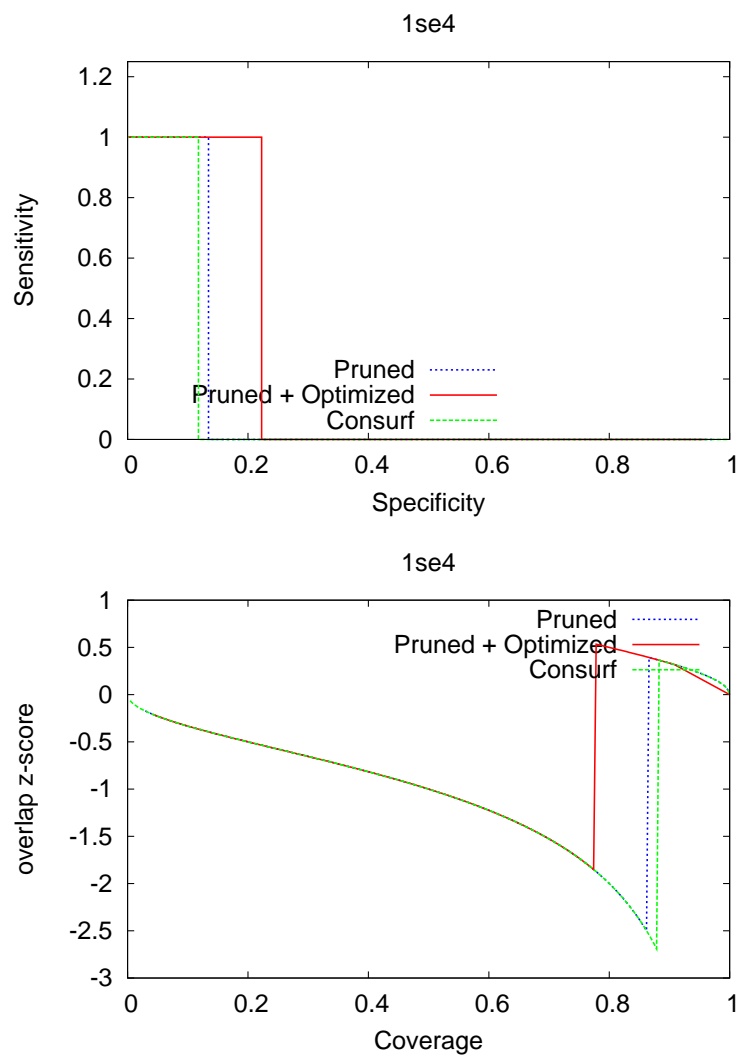
Figure 149: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1rnl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
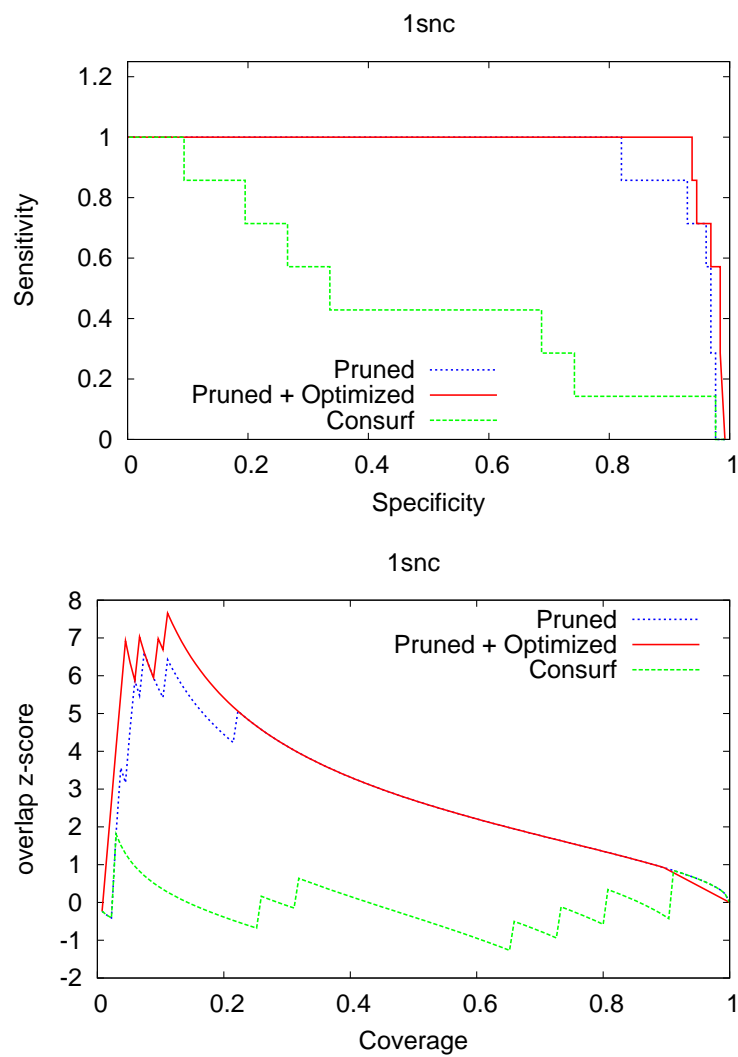
Figure 150: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1se4. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
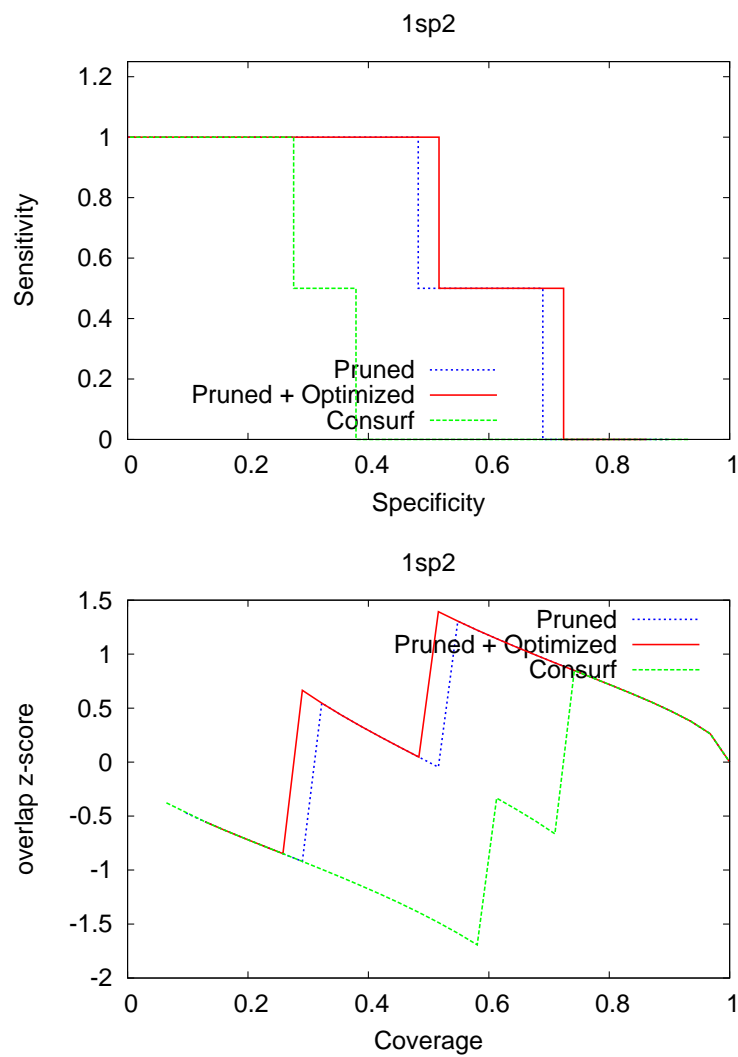
Figure 151: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1snc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
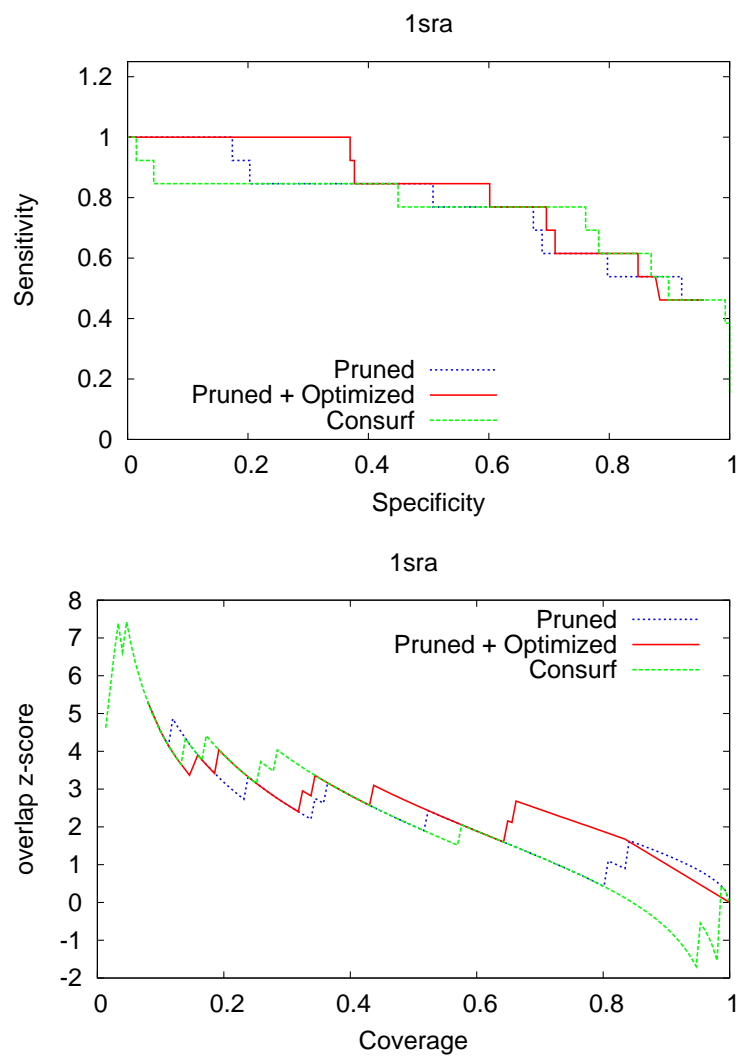
Figure 152: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1sp2. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
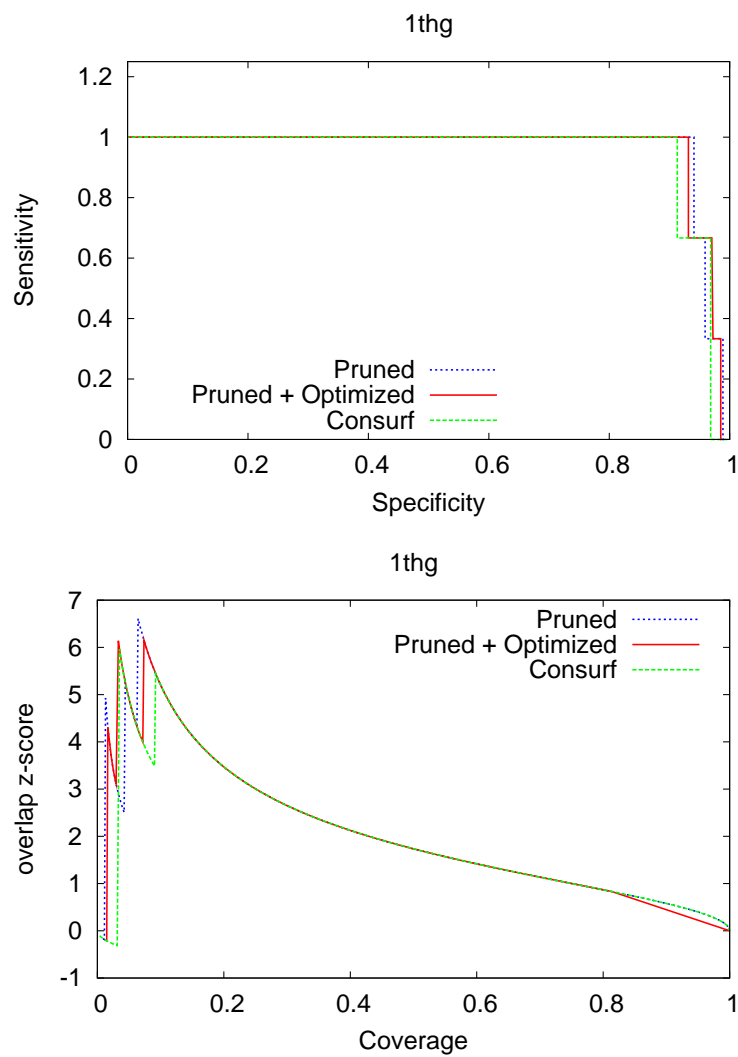
Figure 153: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1sra. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
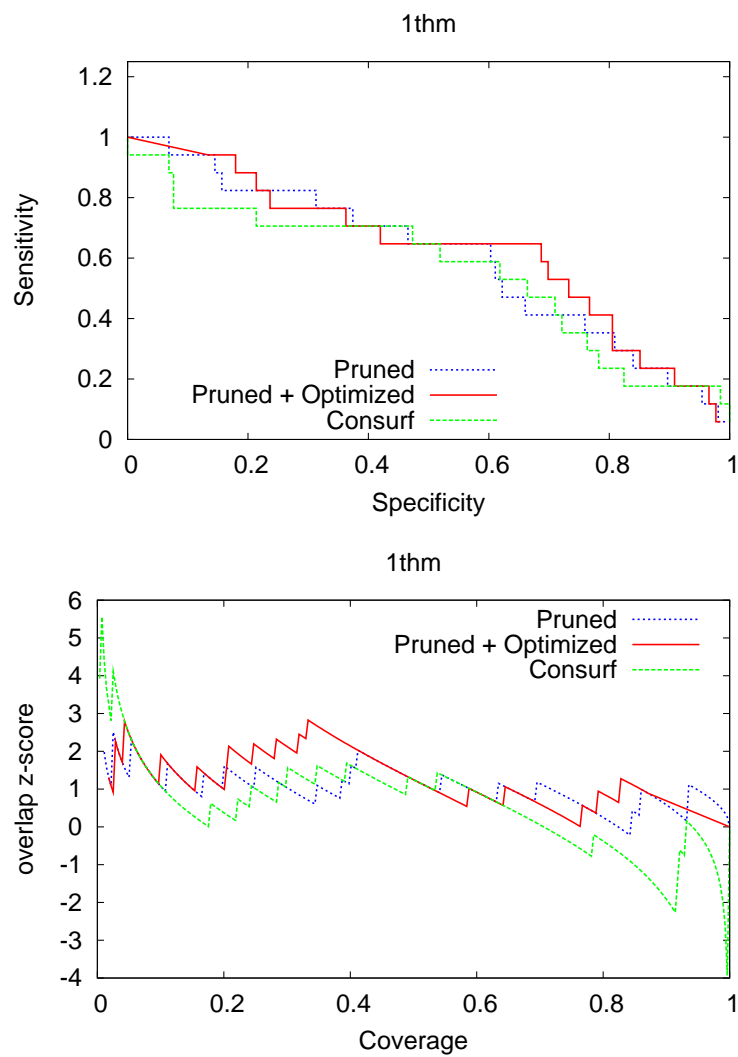
Figure 154: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1thg. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
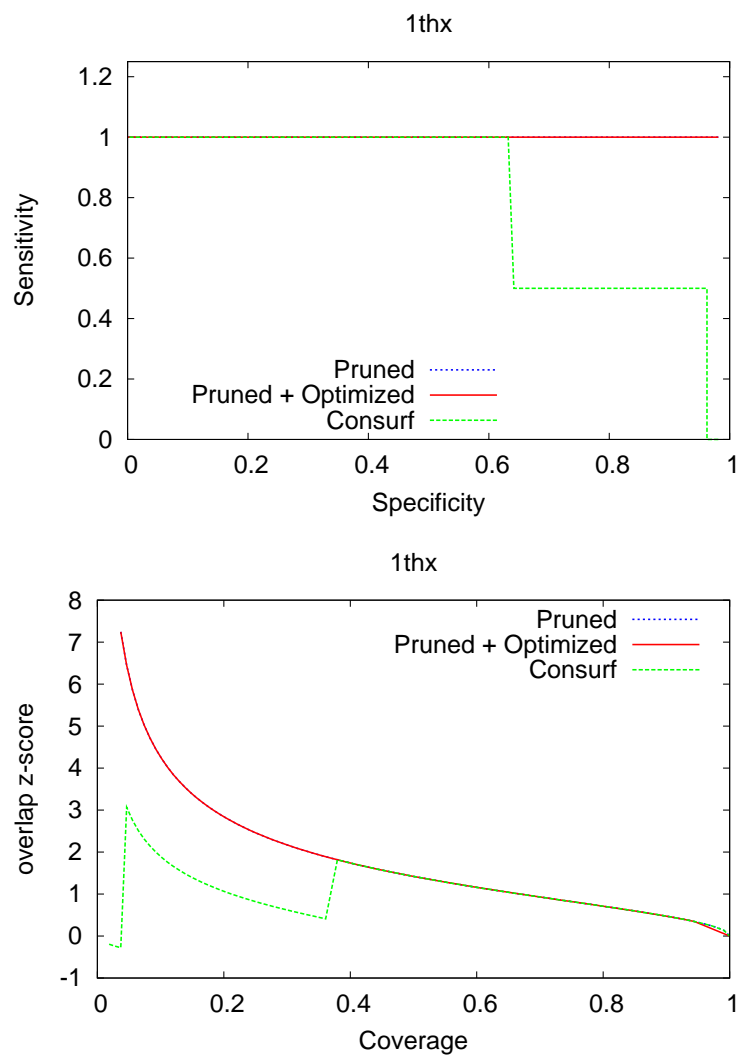
Figure 155: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1thm. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
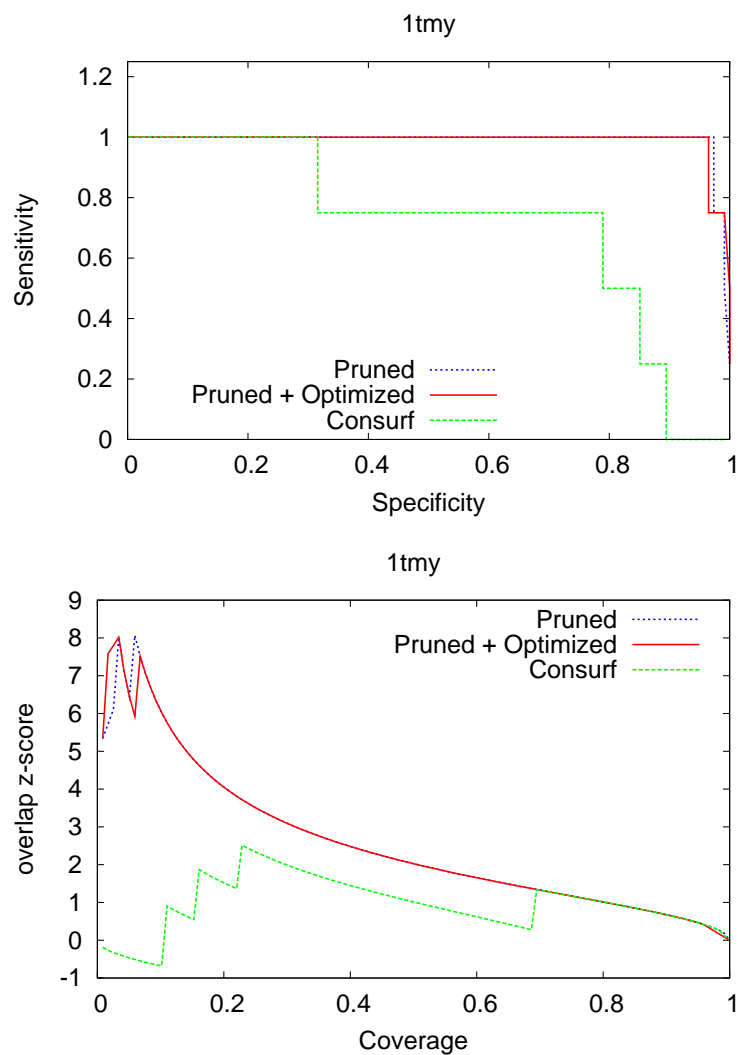
Figure 156: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1thx. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
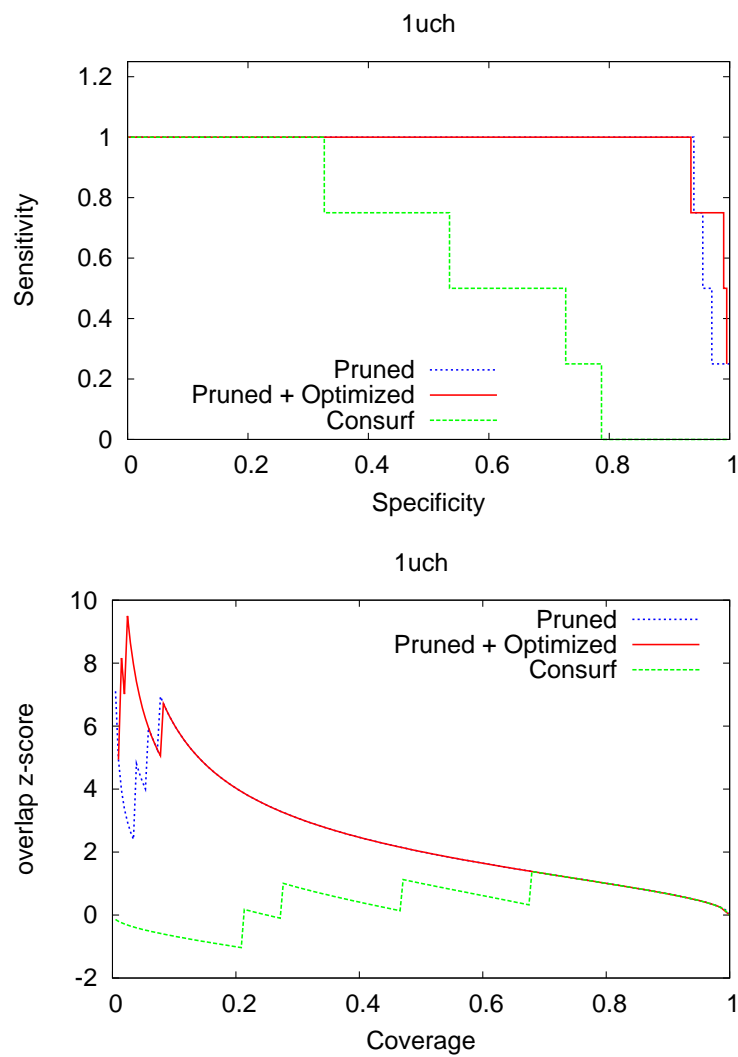
Figure 157: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1tmy. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
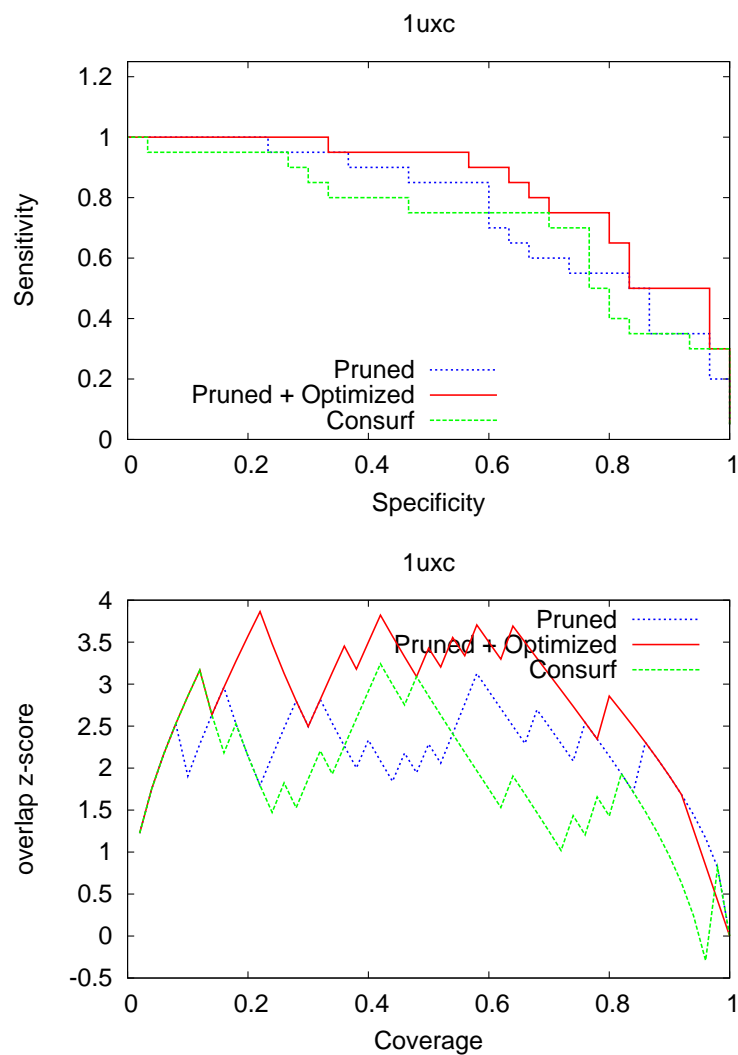
Figure 158: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1uch. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
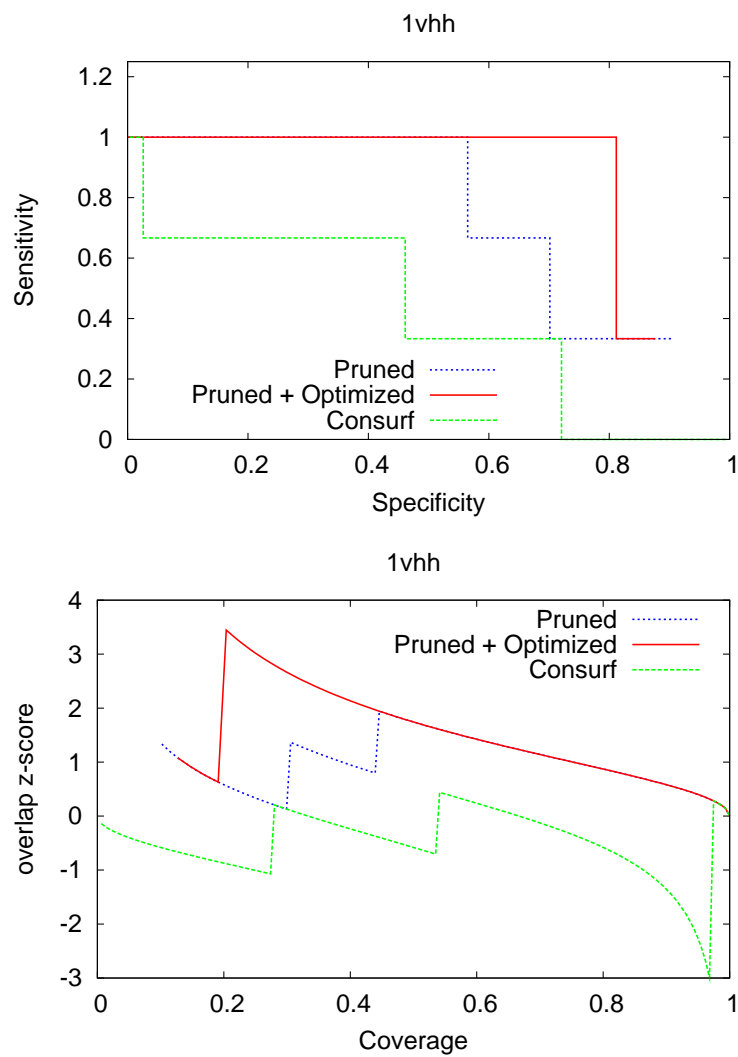
Figure 159: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1uxc. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
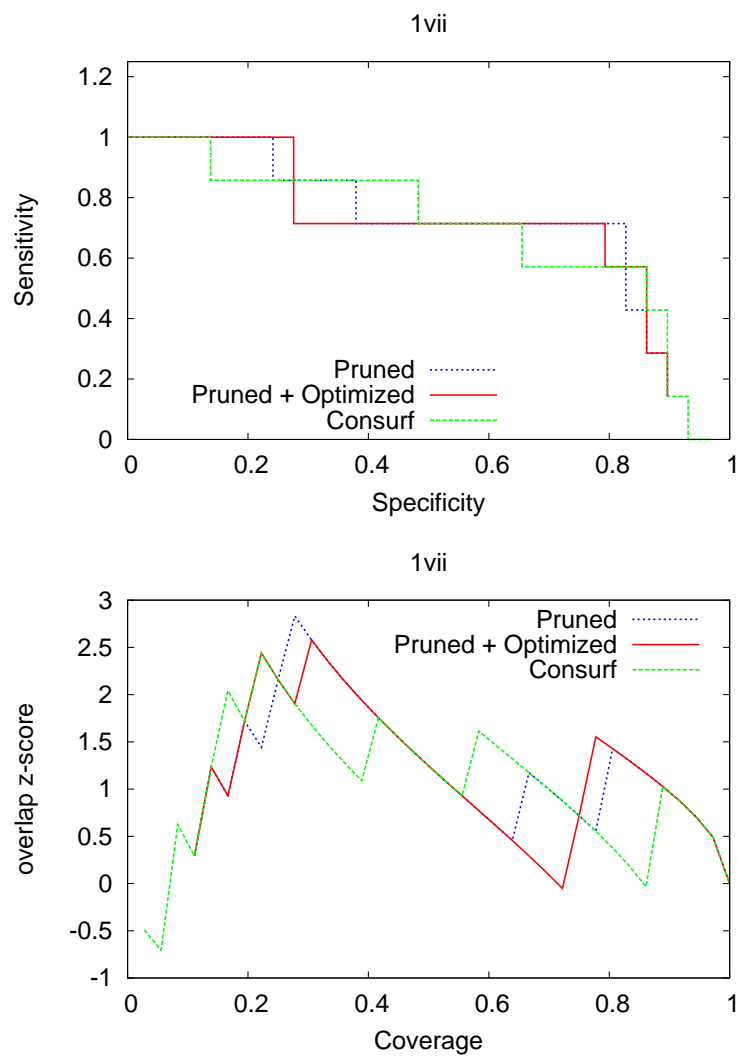
Figure 160: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1vhh. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
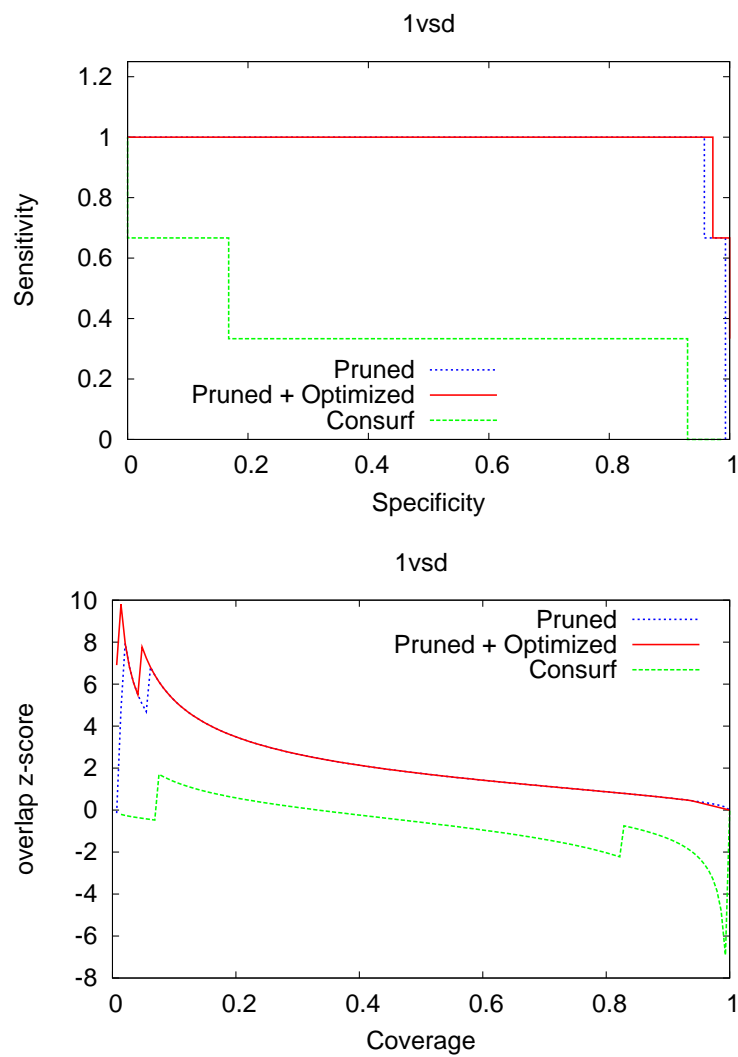
Figure 161: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1vii. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
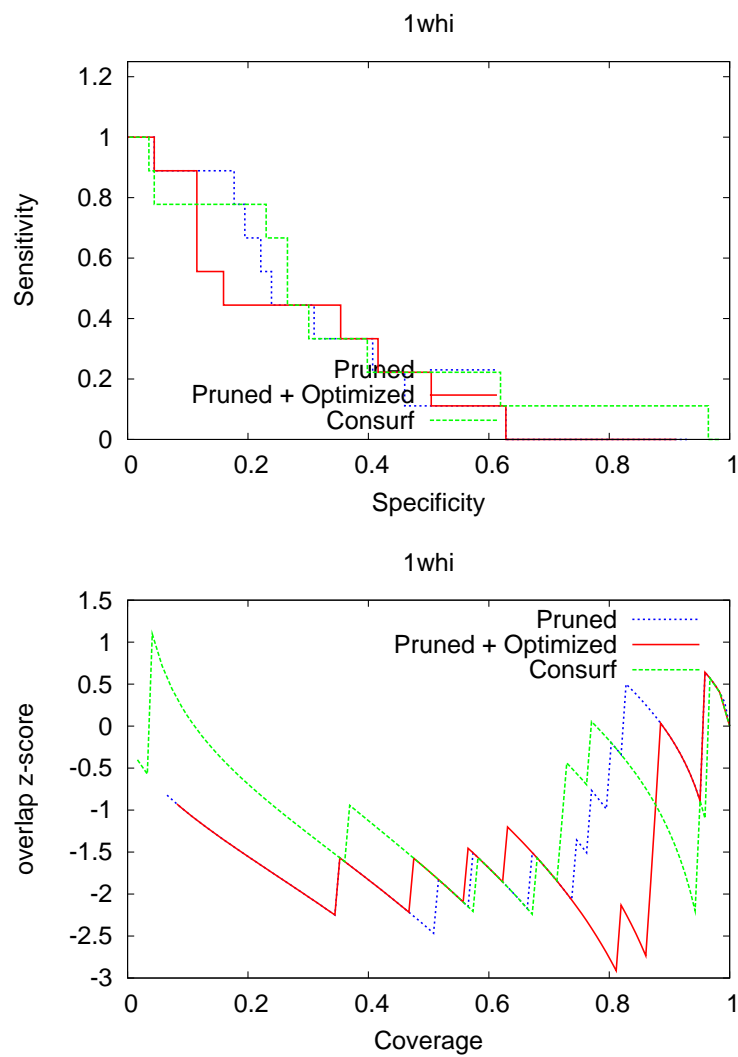
Figure 162: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1vsd. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
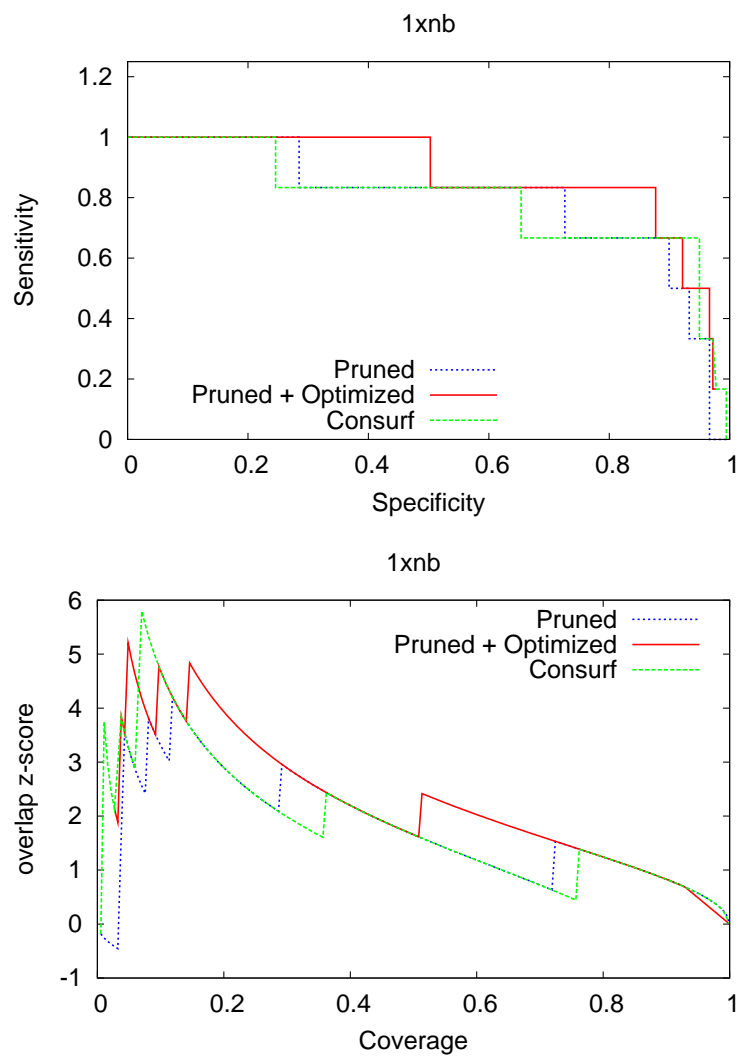
Figure 163: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1whi. b) The overlap z-scores as a function of rank coverage for the three methods is shown.

Figure 164: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 1xnb. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
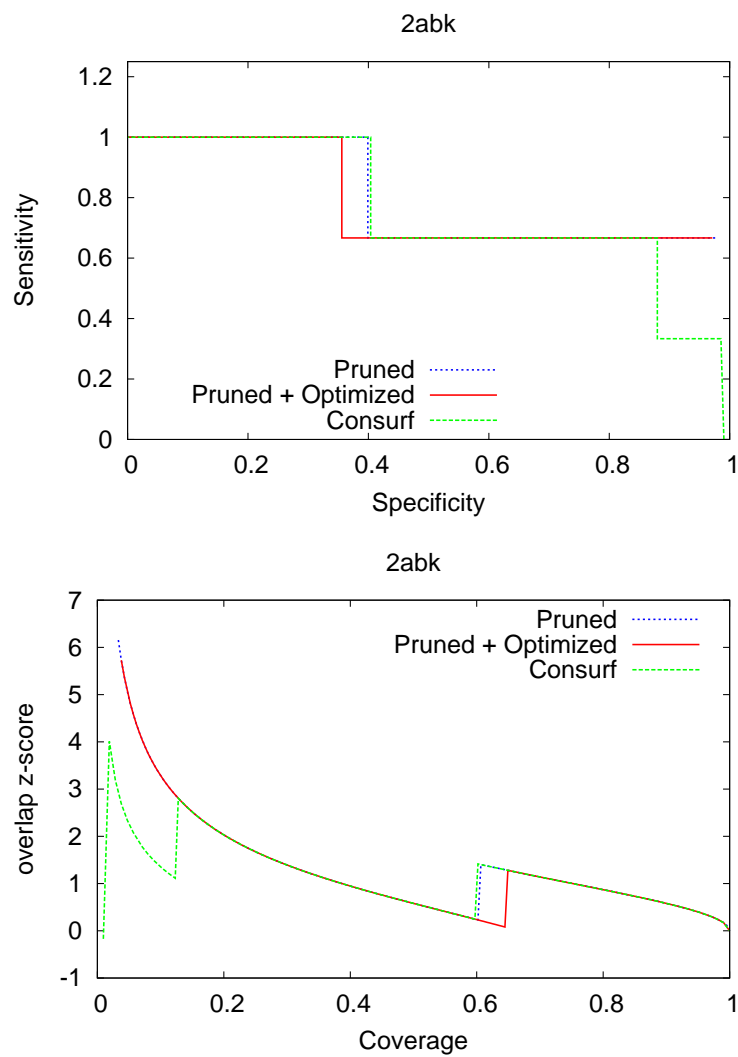
Figure 165: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2abk. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
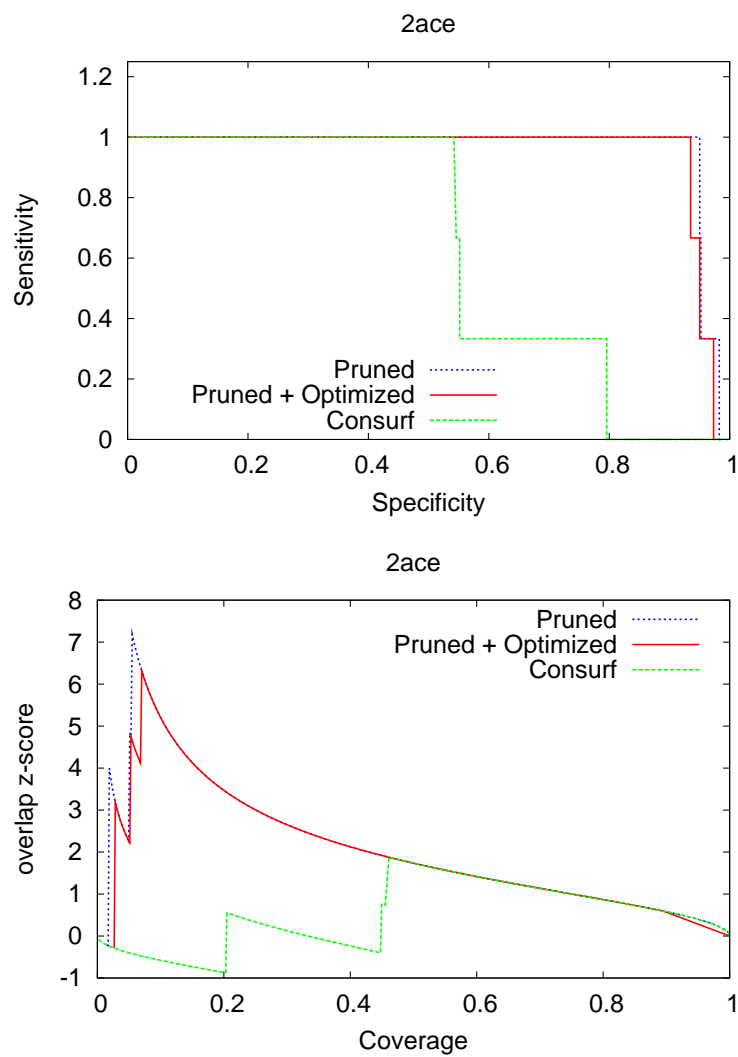
Figure 166: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2ace. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
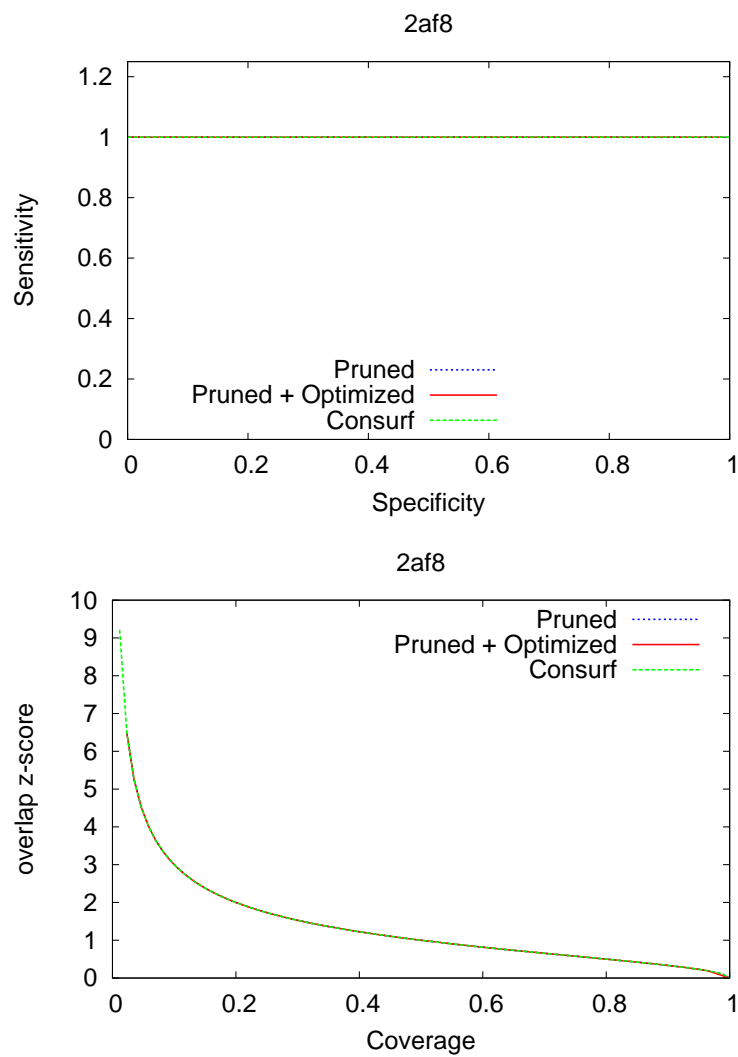
Figure 167: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2af8. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
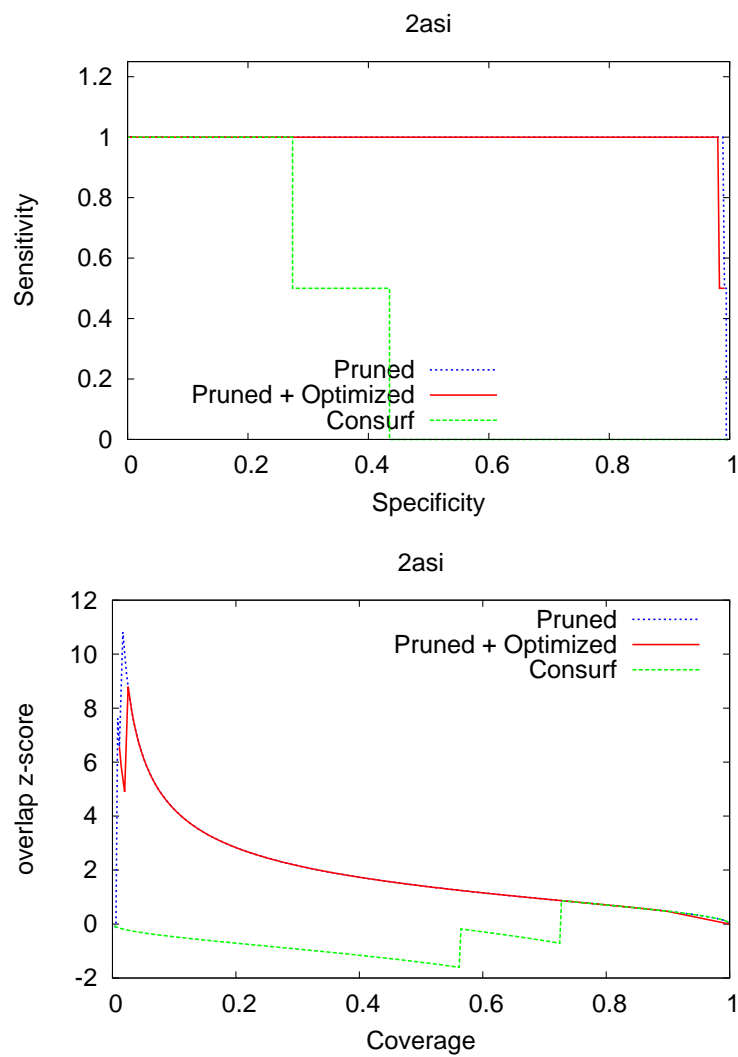
Figure 168: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2asi. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
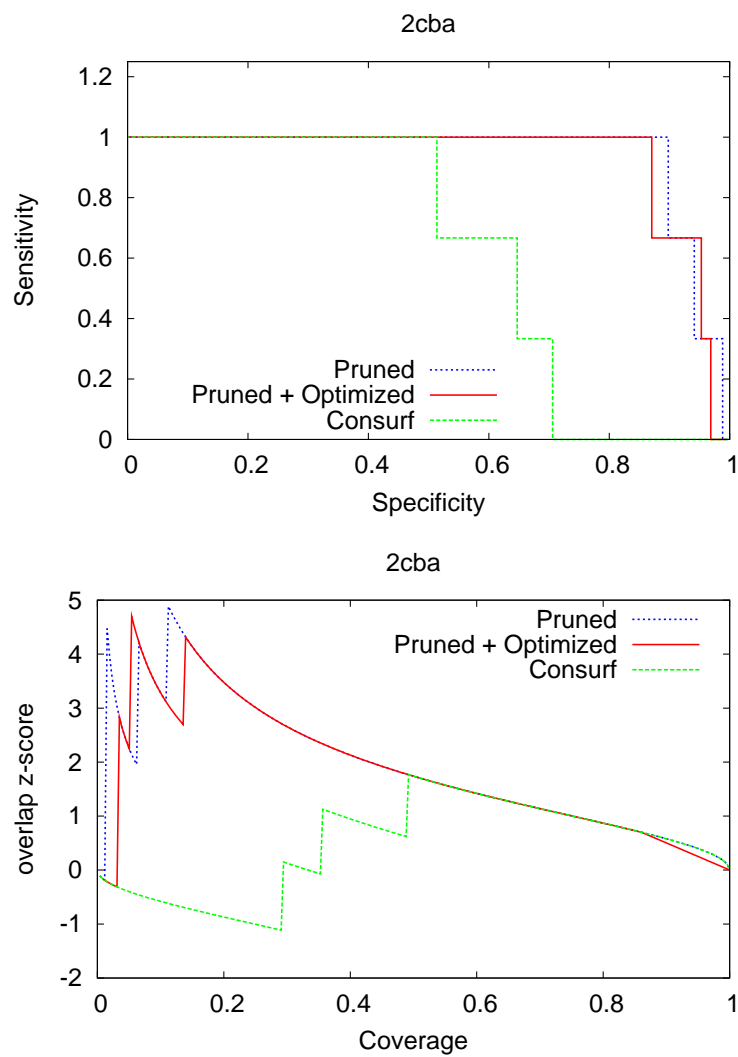
Figure 169: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2cba. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
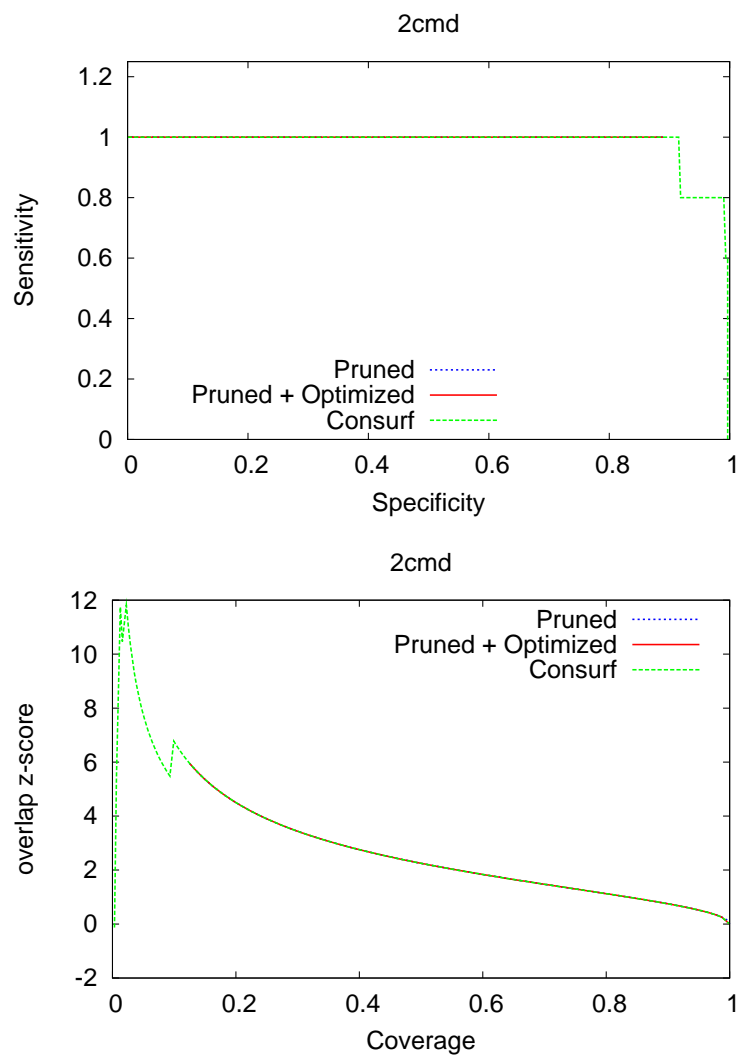
Figure 170: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2cmd. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
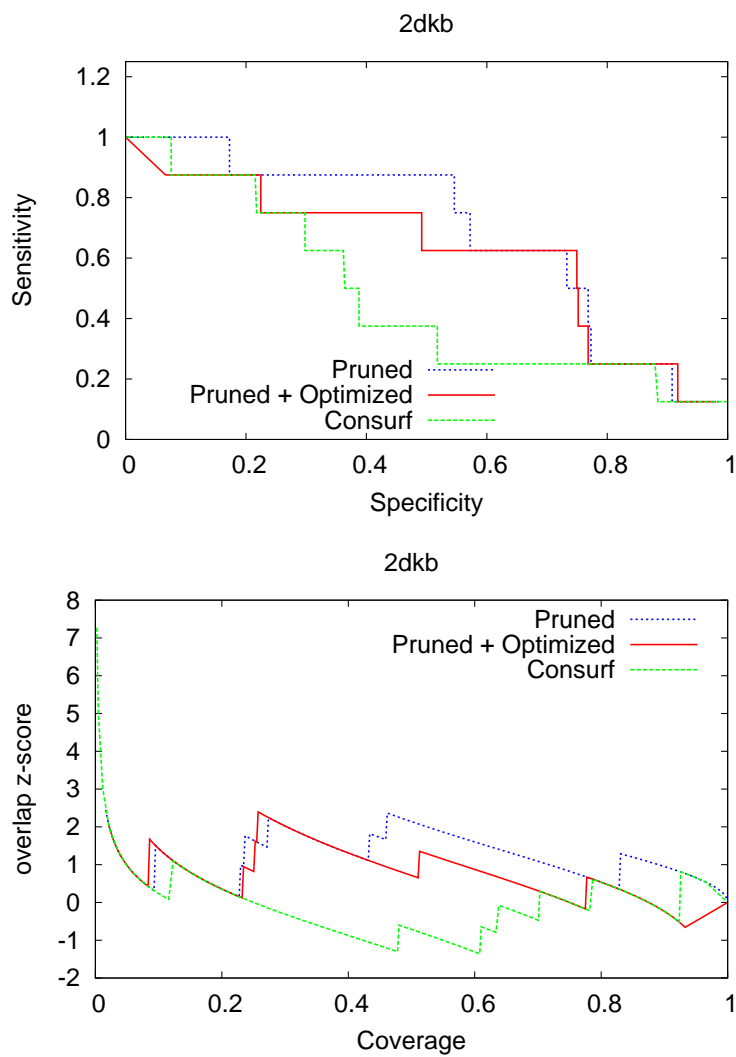
Figure 171: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2dkb. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
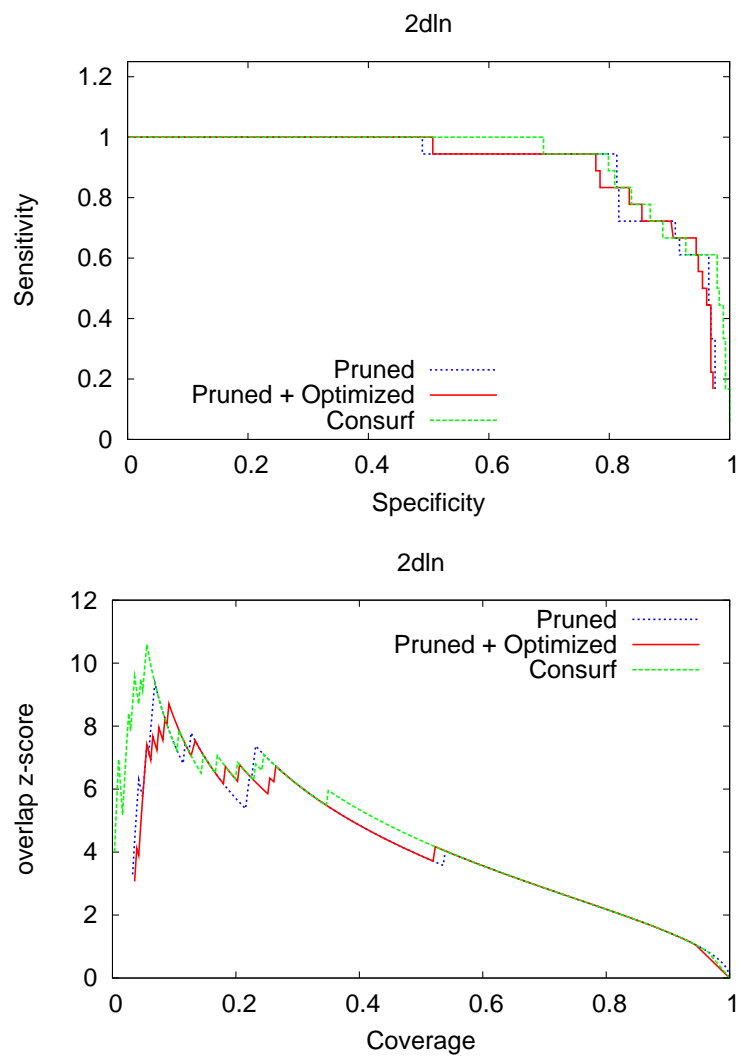
Figure 172: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2dln. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
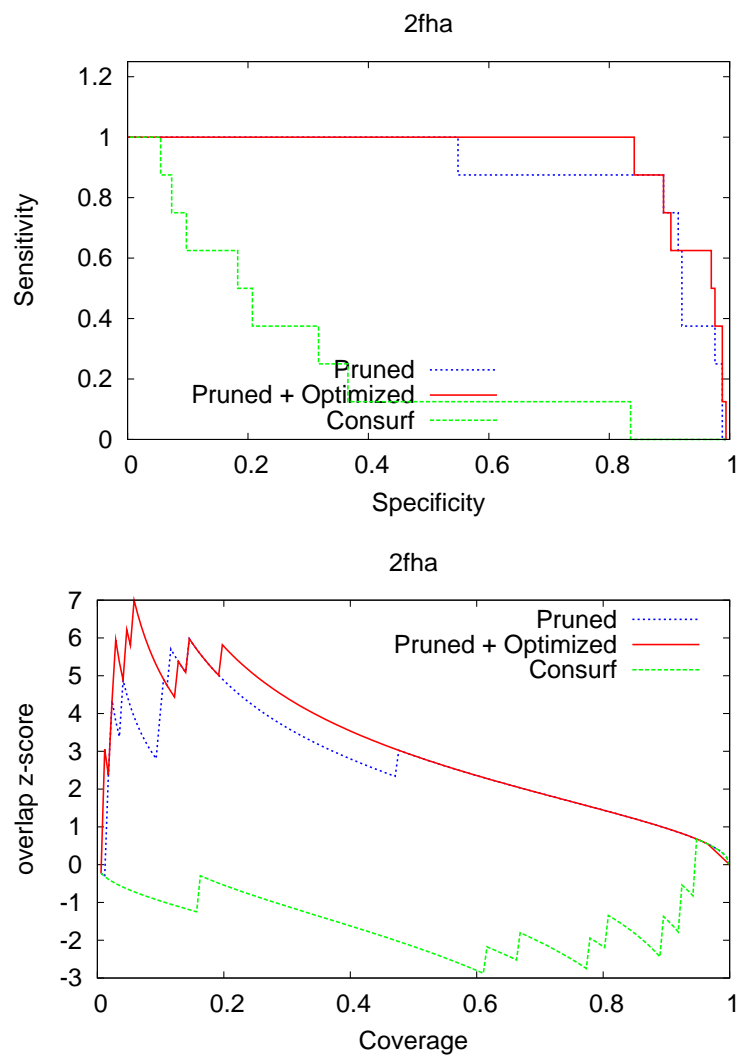
Figure 173: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2fha. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
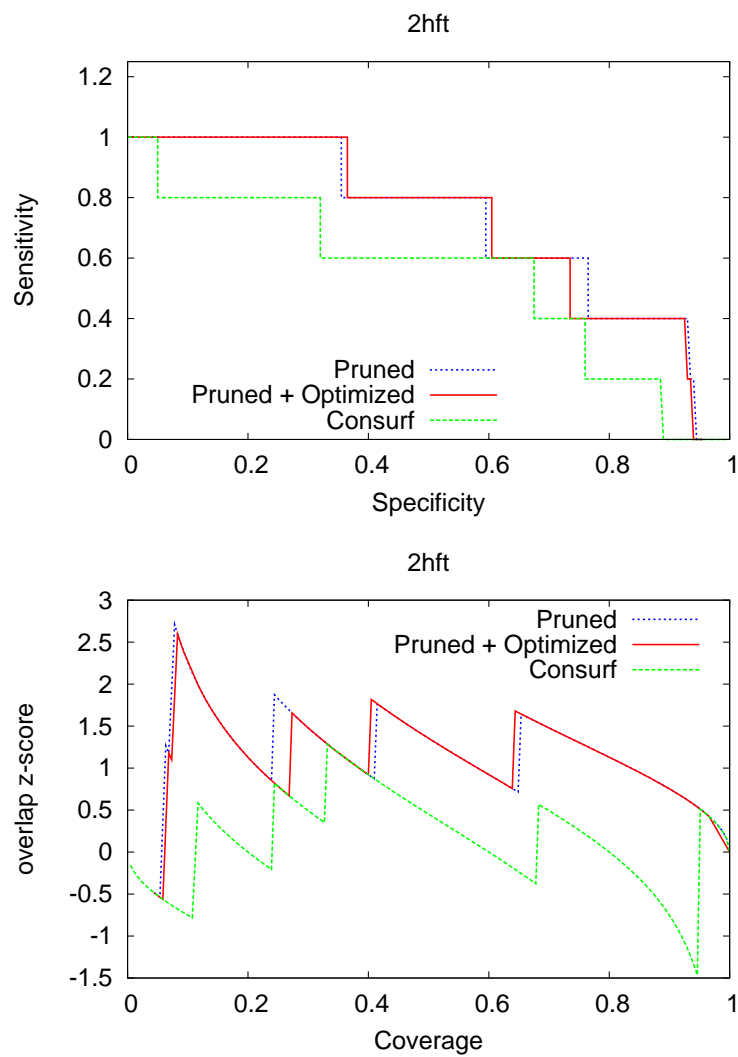
Figure 174: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2hft. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
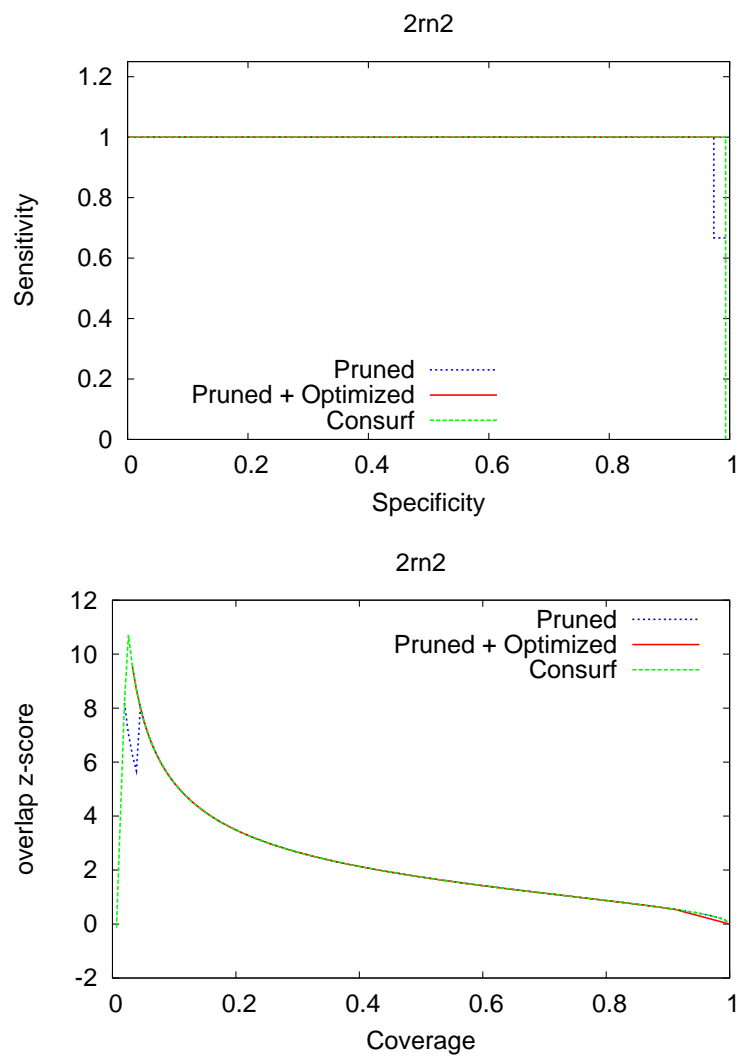
Figure 175: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2rn2. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
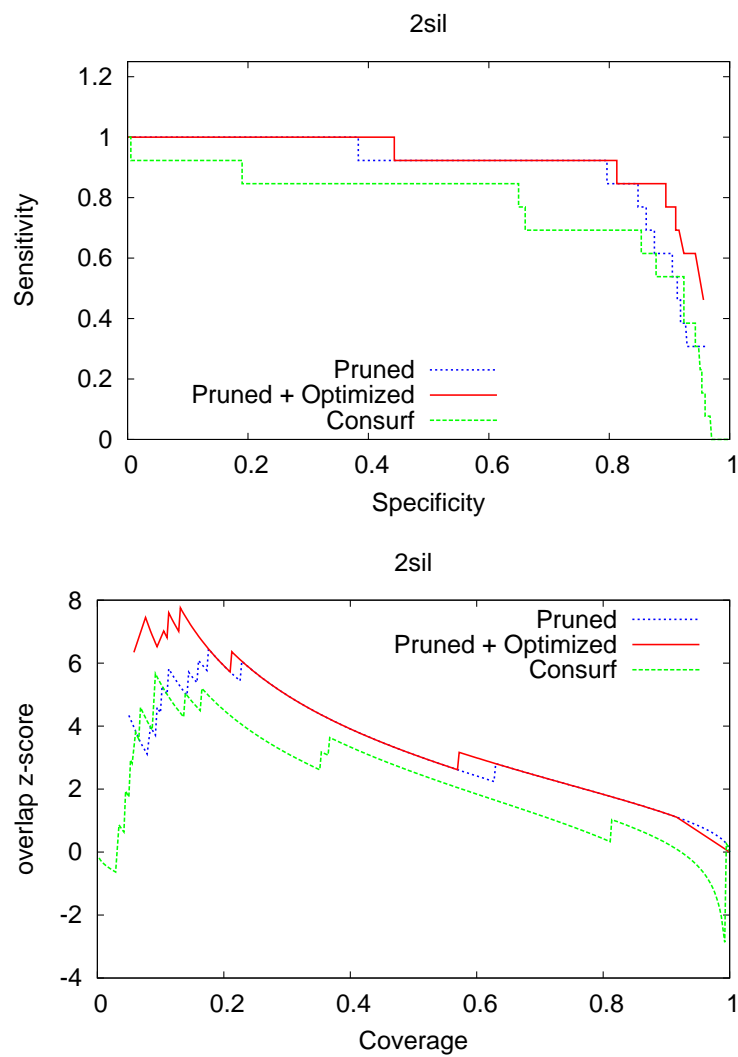
Figure 176: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2sil. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
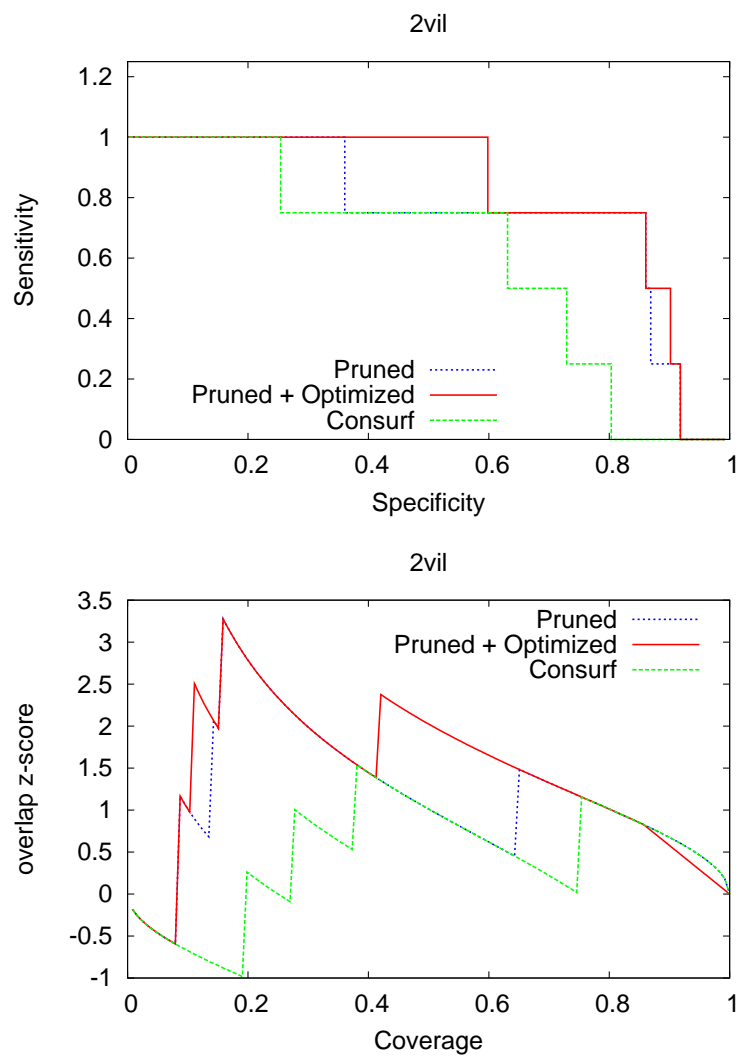
Figure 177: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 2vil. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
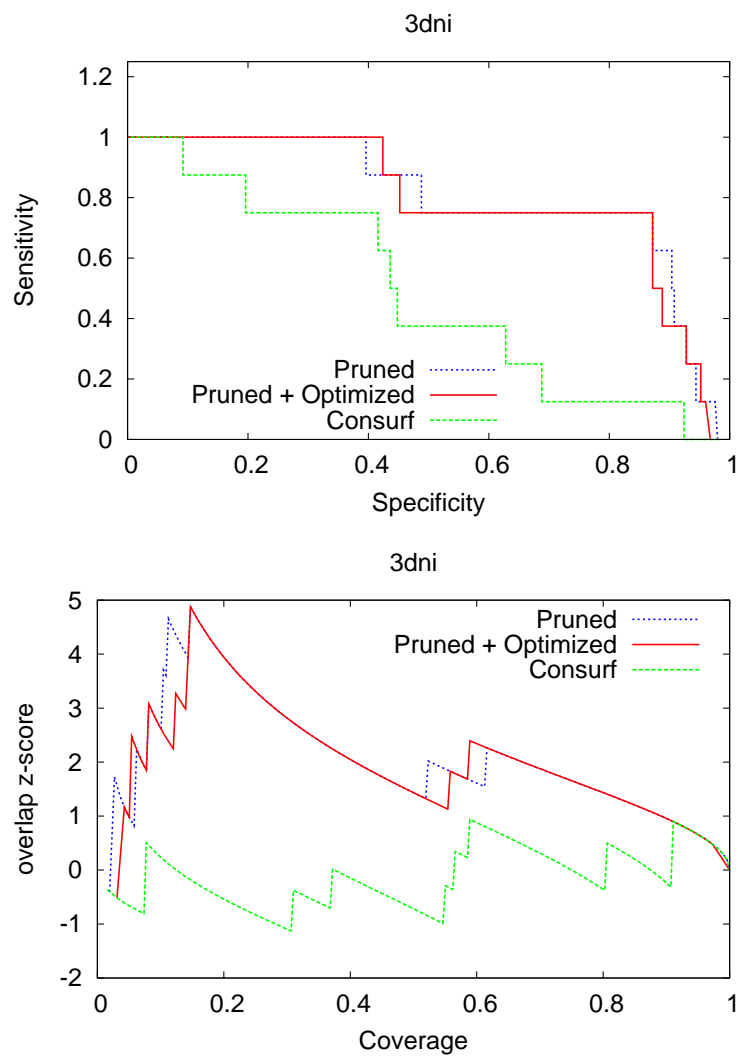
Figure 178: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 3dni. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
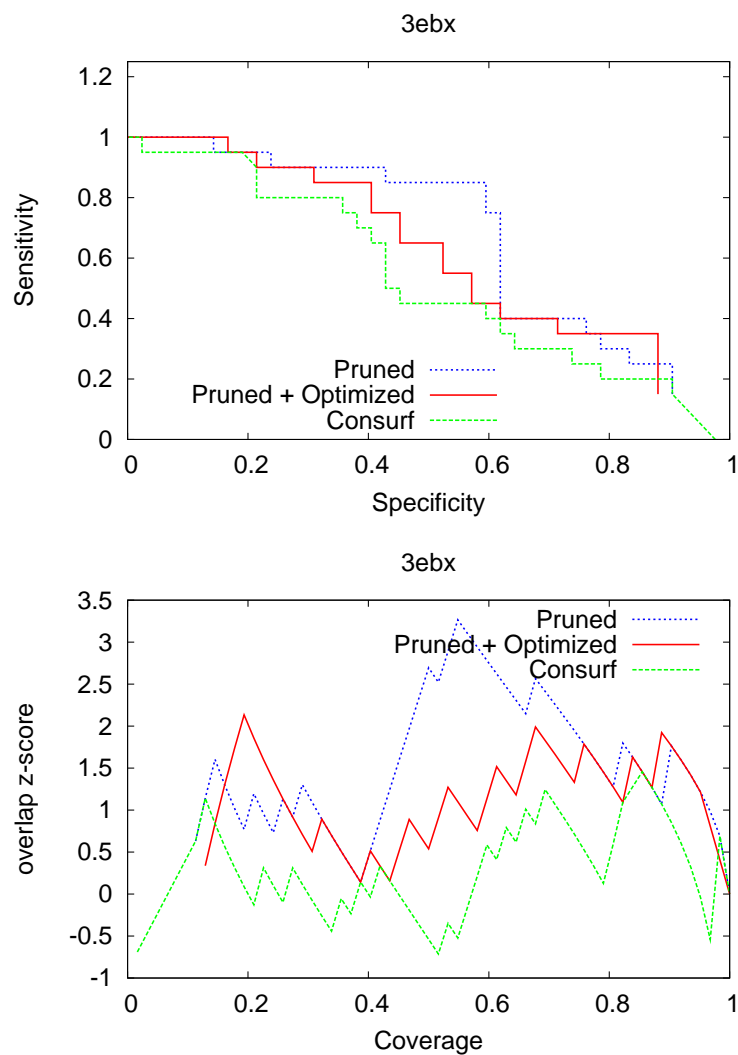
Figure 179: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 3ebx. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
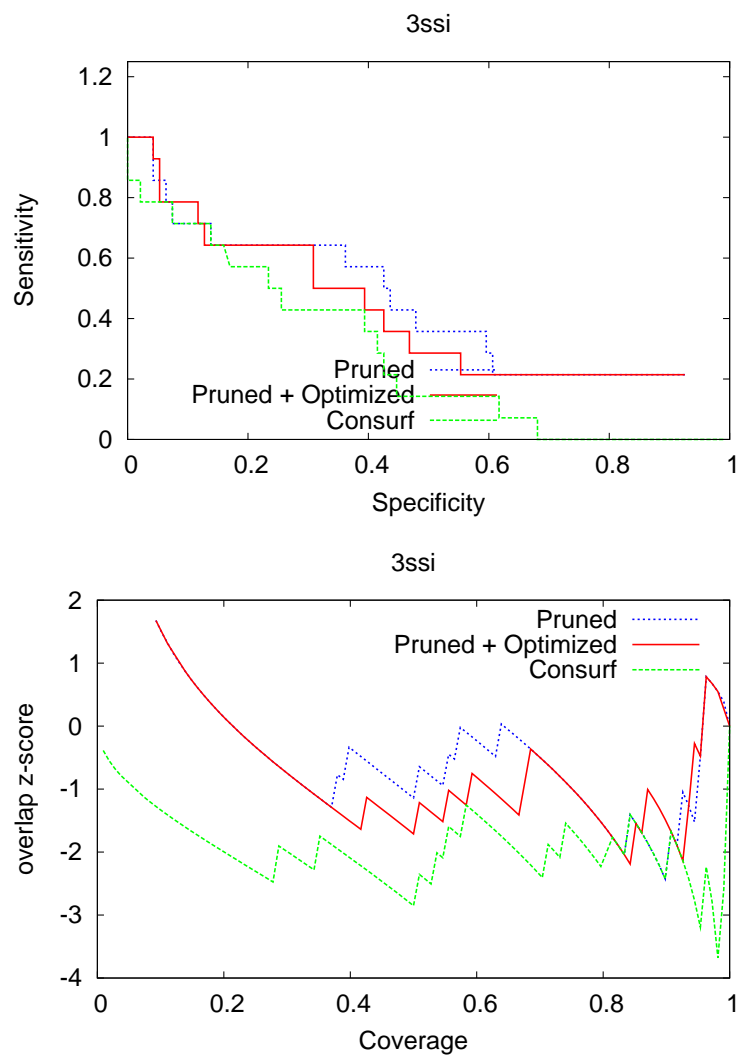
Figure 180: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 3ssi. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
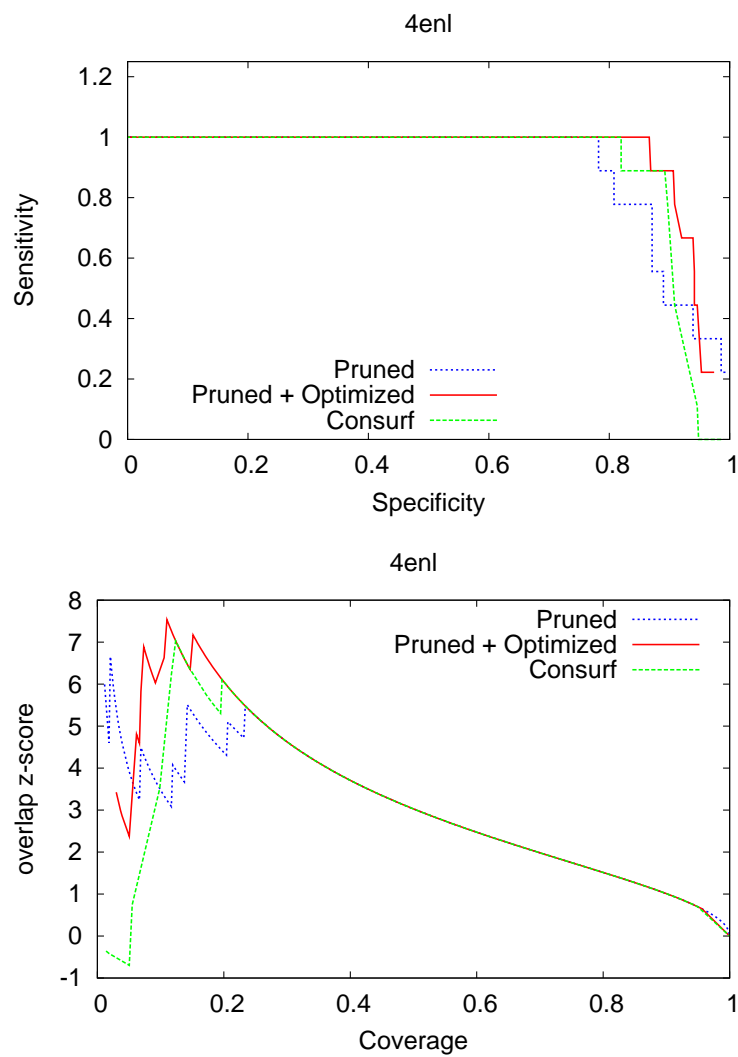
Figure 181: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 4enl. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
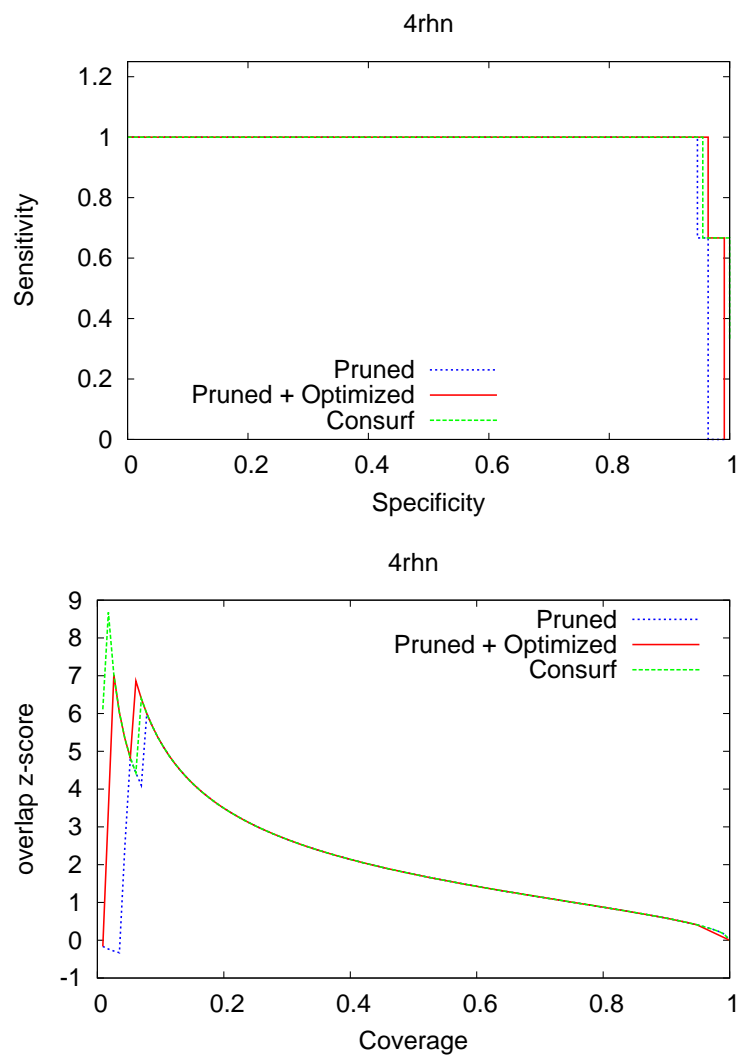
Figure 182: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 4rhn. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
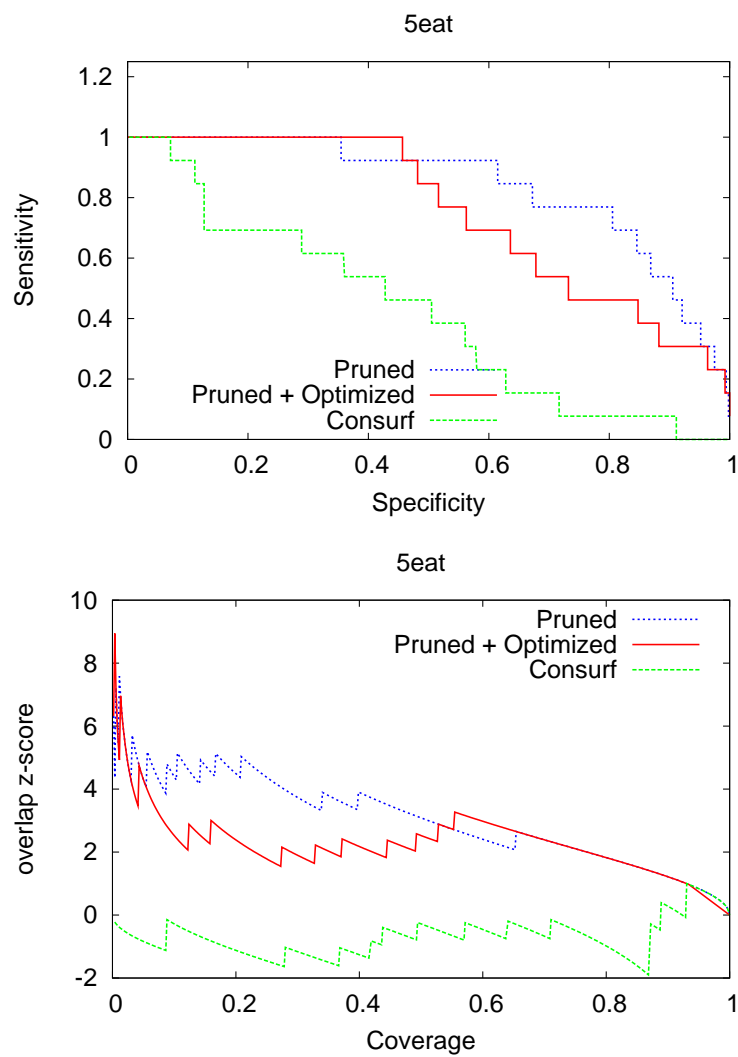
Figure 183: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 5eat. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
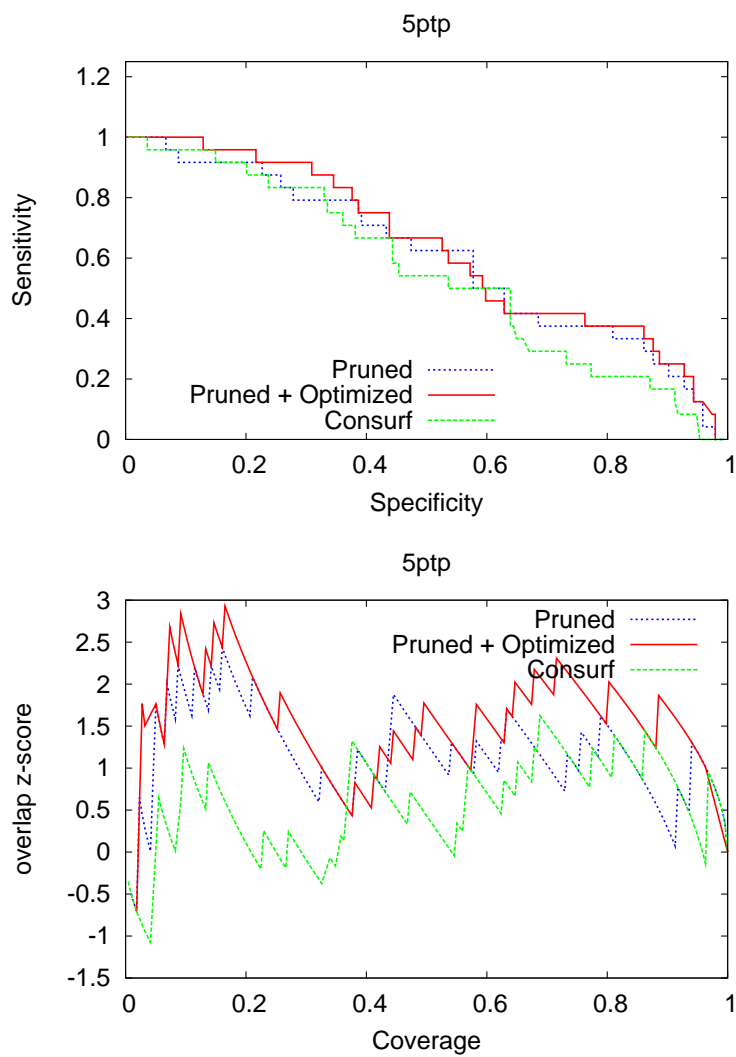
Figure 184: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 5ptp. b) The overlap z-scores as a function of rank coverage for the three methods is shown.
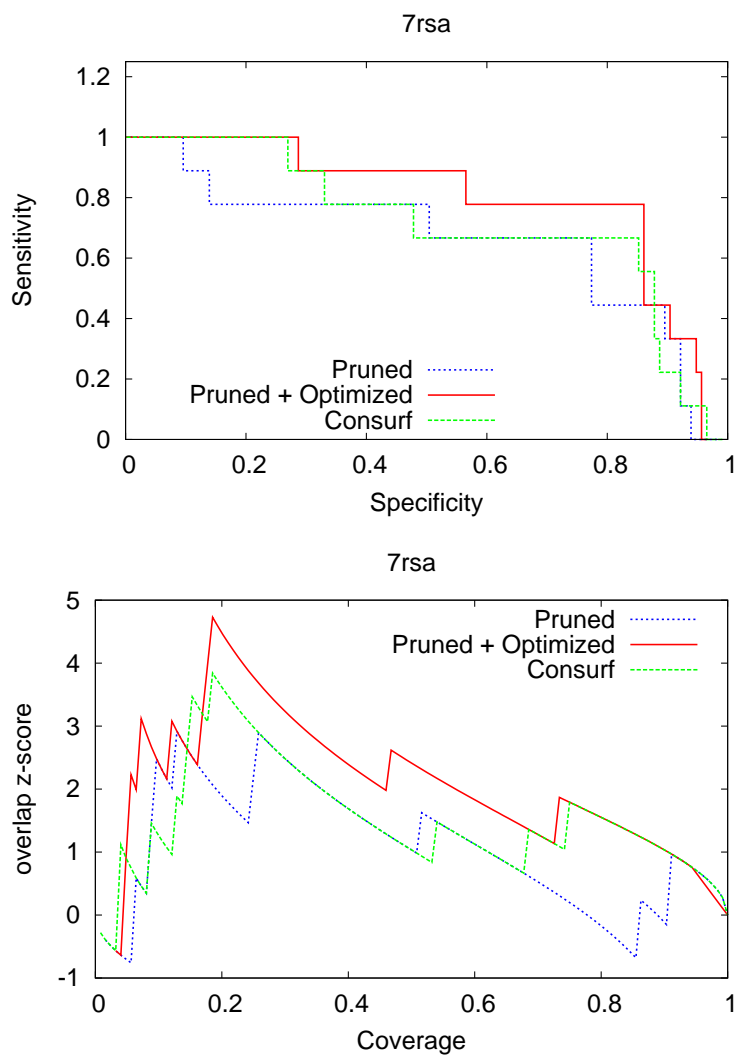
Figure 185: a) The sensitivity and specifity is shown for the pruned, pruned + optimized and the comparison method Consurf for PDBID 7rsa. b) The overlap z-scores as a function of rank coverage for the three methods is shown.