# Table S1: Comparison of 3D models with Quality Measures (Post-prediction analysis)

| Allergen Name | Protein | Sequence Identity to template (%) | Structural RMSD (Å) | Coverage (%) | QMEAN[#] Score (QMEAN Z-score) | ProSA[%] (Z-score) | Verify_3D[&] Average Score > 0.2 (%) |
|---|---|---|---|---|---|---|---|
| Hev b 2 | rubber (latex), 1,3-glucanase | 45.9 | 1.3<br>1.2* | 95.5 (300/314)<br>98.4 (309/314)* | 0.777 (0.03) | -8.60 | 99.68<br>(Pass) |
| Gly m 4 | soybean, stress-induced protein SAM22 | 44.9 | 2.3<br>2.7* | 98.1 (153/156)<br>100.0 (156/156)* | 0.739 (-0.15) | -7.13 | 100<br>(Pass) |
| Sol i 3 | fire ant, venom allergen III | 42.8 | 1.5<br>1.7* | 92.6 (199/215)<br>96.3 (207/215)* | 0.652 (-1.25) | -5.61 | 66.98<br>(Warning) |
| Ara h 1 | peanut, vicilin | 38.7 | 1.5<br>1.6* | 79.8 (332/416)<br>83.9 (349/416)* | 0.695 (-0.91) | -5.93 | 88.42<br>(Pass) |
| Dau c 1 | carrot, pathogenesis-related protein, PR-10 | 37.9 | 1.3<br>1.3* | 95.4 (146/153)<br>99.3 (152/153)* | 0.725 (-0.27) | -7.00 | 85.06<br>(Pass) |
| Der f 1.0106 | American house dust mite, cysteine protease | 31.6 | 1.9<br>1.9* | 90.1 (191/212)<br>99.1 (210/212)* | 0.686 (-0.89) | -6.25 | 67.61<br>(Warning) |
| Can f 2 | salivary canine allergen, lipocalin | 26.8 | 1.5<br>1.6* | 89.8 (141/157)<br>94.3 (148/157)* | 0.633 (-1.28) | -4.30 | 94.30<br>(Pass) |
| Ara h 2 | peanut, conglutin | 22.4 | 3.1<br>2.2* | 41.0 (55/134)<br>63.4 (85/134)* | 0.452 (-3.15) | -4.00 | 66.14<br>(Warning) |
| Bla g 4 | German cockroach, calycin | 15.4 | 3.0<br>3.0* | 35.2 (64/182)<br>77.0 (141/182)* | 0.307 (-4.68) | -2.67 | 55.93<br>(Fail) |
| Aed a 2 | mosquito, salivary, odorant binding protein, protein D7 | 9.8 | 3.2<br>2.5* | 53.7 (66/123)<br>94.0 (116/123)* | 0.415 (-3.55) | -5.63 | 38.71<br>(Fail) |
| | | 11.3 | 3.2<br>2.9* | 65.3 (81/124)<br>94.0 (116/124)* | 0.416 (-3.50) | -4.03 | 58.40<br>(Fail) |

[#] QMEAN scores were calculated by the QMEAN web server (http://swissmodel.expasy.org/qmean/cgi/index.cgi). High quality 3D models have typically QMEAN Z-scores between -1.0 and 0.5 (mean Z-score = -.65). A lower Z-score is observed for medium quality models (mean Z-score = -1.75), and low quality models have typically Z-scores below -3.0.

[%] The ProSA Z-scores were calculated through the ProSA web server (https://prosa.services.came.sbg.ac.at/prosa.php). The Z-score measures the deviation of a knowledge-based energy value of a model relative to the energy distribution derived from random conformations of proteins. The Z-score of a model has to be compared to the Z-scores of experimental determined X-ray and NMR structures which are size dependent in the assessment of model quality.

[&] The 3D-1D scores from Verify_3D were calculated through the web server at the NIH MBI Laboratory for Structural Genomics and Proteomics at UCLA ( http://nihserver.mbi.ucla.edu/SAVES/). The score measures the compatibility of a 3D protein structure with its sequence by using the statistical preferences of each of the twenty amino acids with the structural environment. We used as a quantitative measure for global model quality the percentage of amino acids in the models with a score greater than 0.2. In general, native like structures tend to have very high ratios near 100%, and 3D models of low quality tend to have low ratios (<60%).

The classification of the models in pass, warning and fail is taken from the output of the Verify_3D web server.