Table 7: Summary of Integrative Structure Determination of A3G_Vif_CRL5 with rigid representation

1) Gathering information	
Prior models	CBFB: comparative model, template 4N9F:F Vif: comparative model, template 4N9F:G EloB: comparative model, template 4N9F:D EloC: comparative model, template 4N9F:E CUL5: comparative model, template 4N9F:C Rbx2: comparative model, template 1LDJ:B A3G: comparative model, template 5K81:E
Physical principles and statistical preferences	Excluded volume Sequence connectivity
Experimental data	132 DSSO chemical cross-links Predicted residue-protein contacts from mutagenesis studies; A3G residues 126-132 and Vif residues 40-45
2) Representing the system	
Composition (number of copies)	A3G: 1 CBFB: 1 CUL5: 1 EloB: 1 EloC: 1 Rbx2: 1 Vif: 1
Atomic (structured) components	A3G: 6-194, 200-243, 258-380 CBFB: 1-156 CUL5: 11-118, 133-302, 308-382, 405-780 EloB: 1-105 EloC: 17-112 Rbx2: 27-113 Vif: 6-154, 166-175
Unstructured components	A3G: 1-5, 195-199, 244-257, 381-384 CBFB: 157-182 CUL5: 1-10, 119-132, 303-307, 383-404 EloB: 106-161 EloC: 1-16 Rbx2: 1-26 Vif: 1-5, 155-165
Resolution of structured components	1 [R1], 10 [R10] residues per bead
Resolution of unstructured components	5 [R5] residues per bead
Structural coverage	89.1 %
Rigid body (RB) definitions	RB1: $CBFB_{1-156}$,Vif_{6-154},$Vif_{166-175}$,$EloB_{1-105}$,$EloC_{17-112}$,$CUL5_{1}$ RB2: $A3G_{6-194}$ RB3: $A3G_{200-243}$,$A3G_{258-380}$
Spatial restraints encoded into scoring function	Excluded volume; applied to the R1 representation Sequence connectivity; applied to the R1 representation Cross-link restraints; applied to the R1 representation Residue-protein proximity restraints; applied to the R1 representation
3) Structural Sampling	
Sampling method	Replica Exchange Gibbs sampling, based on Metropolis Monte Carlo
Replica exchange temperature range	1.0 - 2.5
Number of replicas	8
Number of runs	50
Number of structures generated	3000000
Movers for flexible string of bead	Random translation up to 4.0 Å
CPU time	6 hours on 20 processors
4) Validating the model	
Models selected for validation	
Number of models after equilibration	3000000
Number of models that satisfy the input information	937225
Number of structures in samples A/B	417675/519550
p-value of non-parametric Kolmogorov-Smirnov two-sample test	0.009 (threshold p-value > 0.05)

Kolmogorov-Smirnov two-sample test statistic, D	0.57
Thoroughness of the structural sampling	
Sampling precision	11.73 Å
Homogeneity of proportions χ^2 test (p-value)/Cramers V value	0.449/0.017 (thresholds: p-value>0.05 OR Cramer's V<0.1)
Number of clusters	2
Cluster populations	cluster 1 : 91.5 % cluster 2 : 5.2 %
Cluster precisions	cluster 1 : 8.38 Å cluster 2 : 9.27 Å
Average cross-correlation between localization probability densities of samples A and B	cluster 1: 0.96 cluster 2: 0.96
Validation by information used for modeling	
Percent of sequence connectivity restraints satisfied per structure	99 %
Percent cross-link restraints satisfied by ensemble	89 %
Percent of residue-protein proximity restraints satisfied by ensemble	98 %
Percent of excluded volume restraints satisfied per structure	99 %
5) Benchmark	
6) Software and data availability	
Software	
Modeling programs	IMP PMI module, version develop-af393bce43 Integrative Modeling Platform (IMP), version develop-af393bce43 MODELLER, version 9.20 MODELLER, version 9.19
Modeling scripts	https://github.com/integrativemodeling/A3G_Vif_CRL5
Homology detection and structure prediction	HHPred, version 2.0.16
Visualization and plotting	UCSF Chimera, version 1.10 Matplotlib, version 3.0.3
Data	
PDB-dev accesion code	TBD