Characterization of an A3G-Vif$_{HIV-1}$-CRL5-CBFβ Structure Using a Cross-linking Mass Spectrometry Pipeline for Integrative Modeling of Host–Pathogen Complexes

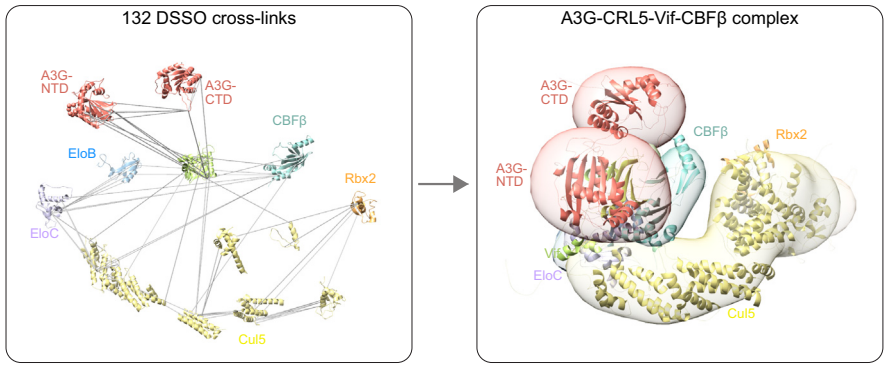
Authors

Robyn M. Kaake, Ignacia Echeverria, Seung Joong Kim, John Von Dollen, Nicholas M. Chesarino, Yuqing Feng, Clinton Yu, Hai Ta, Linda Chelico, Lan Huang, John Gross, Andrej Sali, and Nevan J. Krogan

Correspondence

sali@salilab.org; nevan.krogan@ucsf.edu

Graphical Abstract



In Brief

We present a pipeline that streamlines cross-linking mass spectrometry (XL-MS) data collection, data analysis, and integrative modeling of host–pathogen complexes. Using XL-MS, known atomic structures, and functional genetic data, we determined an integrative structure of the HIV-human A3G-CRL5-Vif-CBFβ complex. This structure illustrates HIV-1 Vif interaction with A3G and captures the structural dynamics and flexibility of the entire A3G-CRL5-Vif-CBFβ complex.

Highlights

- Integrative modeling using cross-links enables modeling of heterogeneous complexes.
- A pipeline that streamlines modeling of host–pathogen complexes is presented.
- The A3G-Vif$_{HIV-1}$-CRL5-CBFβ integrative structure in solution is determined.
- The structure shows how Vif recruits A3G and the complex structural dynamics.

2021, Mol Cell Proteomics *20*, 100132

https://doi.org/10.1016/j.mcpro.2021.100132

Characterization of an A3G-Vif$_{HIV-1}$-CRL5-CBFβ Structure Using a Cross-linking Mass Spectrometry Pipeline for Integrative Modeling of Host–Pathogen Complexes

Robyn M. Kaake[1,2,3,‡], Ignacia Echeverria[1,4,‡], Seung Joong Kim[4], John Von Dollen[1,2], Nicholas M. Chesarino[5], Yuqing Feng[6], Clinton Yu[7], Hai Ta[8], Linda Chelico[6], Lan Huang[7], John Gross[2,8], Andrej Sali[2,4,8,*], and Nevan J. Krogan[1,2,3,*]

Structural analysis of host–pathogen protein complexes remains challenging, largely due to their structural heterogeneity. Here, we describe a pipeline for the structural characterization of these complexes using integrative structure modeling based on chemical cross-links and residue–protein contacts inferred from mutagenesis studies. We used this approach on the HIV-1 Vif protein bound to restriction factor APOBEC3G (A3G), the Cullin-5 E3 ring ligase (CRL5), and the cellular transcription factor Core Binding Factor Beta (CBFβ) to determine the structure of the (A3G-Vif-CRL5-CBFβ) complex. Using the MS-cleavable DSSO cross-linker to obtain a set of 132 cross-links within this reconstituted complex along with the atomic structures of the subunits and mutagenesis data, we computed an integrative structure model of the heptameric A3G-Vif-CRL5-CBFβ complex. The structure, which was validated using a series of tests, reveals that A3G is bound to Vif mostly through its N-terminal domain. Moreover, the model ensemble quantifies the dynamic heterogeneity of the A3G C-terminal domain and Cul5 positions. Finally, the model was used to rationalize previous structural, mutagenesis and functional data not used for modeling, including information related to the A3G-bound and unbound structures as well as mapping functional mutations to the A3G-Vif interface. The experimental and computational approach described here is generally applicable to other challenging host–pathogen protein complexes.

Structural heterogeneity plays a crucial role in host–pathogen interactions. Several pathogens exploit the intrinsic disorder of some of their proteins and the structural plasticity of the host proteins to their advantage during infection (1–3). As a result, host–pathogen protein assemblies are often refractory to traditional structural biology techniques (*i.e.*, X-ray crystallography, NMR spectroscopy, and single particle cryo-electron microscopy). Thus, an alternative approach is needed to solve the structures of these structurally heterogeneous assemblies. Integrative structure modeling, which is based on combining multiple types of input information, is one such approach (4–8). Here, we have devised a pipeline that streamlines structure characterization of host–pathogen complexes by using integrative structure modeling based on data from cross-linking mass-spectrometry (XL-MS) and residue–protein contacts inferred from mutagenesis studies. We present the application of this pipeline to the structure determination of the A3G-bound Vif-CRL5-CBFβ complex.

XL-MS is a cutting-edge experimental technique that can be used to identify interacting proteins and probe interaction

From the [1]Department of Cellular and Molecular Pharmacology, California Institute for Quantitative Biosciences, [2]Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, California, USA; [3]Gladstone Institute of Data Science and Biotechnology, J. David Gladstone Institutes, San Francisco, California, USA; [4]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA; [5]Divisions of Human Biology and Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA; [6]Department of Biochemistry, Microbiology, Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; [7]Department of Physiology & Biophysics, University of California, Irvine, California, USA; [8]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California, USA

[‡]These authors contributed equally to this work.

[*]For correspondence: Nevan J. Krogan, nevan.krogan@ucsf.edu; Andrej Sali, sali@salilab.org.

Present address for Seung Joong Kim: Department of Physics and Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea.

Present address for Yuqing Feng: Department of Immunology, 1 King's College Cir. MSB 7302, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

https://doi.org/10.1016/j.mcpro.2021.100132

interfaces (9–14). Chemical cross-linkers that covalently bridge proximal reactive residues are identified using high-resolution MS analysis. Compared with traditional atomic resolution structure determination approaches, XL-MS requires a lower amount of protein, allows for versatile buffer conditions, is relatively robust to protein impurities, and can be applied to compositionally and structurally heterogeneous protein assemblies. One of the major challenges facing XL-MS techniques is the complexity of cross-linked peptide spectra (15), which leads to a large peptide search space for analysis programs. However, the development of MS-cleavable cross-linkers (16–18) and improved speed and sensitivity of MS^3 protocols, along with improved computational analyses (14), have made the identification of high-confidence cross-linked peptides more straightforward. Despite these advances, estimating the confidence of each unique cross-link remains challenging. To this end, we developed methods for automatically deconvolving ambiguous spectra assignments as well as methods for quantifying, scoring, filtering, and visualizing cross-linking data, which in turn facilitate use of cross-links in integrative structure modeling.

Integrative structure modeling benefits from multiple types of experimental data to maximize the accuracy, precision, completeness, and efficiency of structure characterization (7, 19, 20). In addition to XL-MS data, the integrative pipeline described here also uses data from mutagenesis studies. For example, mutagenesis studies coupled with two-hybrid interaction mapping techniques (21, 22) can often be used to identify protein–protein interface residues, although they do not provide information about the positions and orientations of the interacting proteins. We developed a restraint on the proximity between residues predicted to be at protein–protein interfaces (PPI), based on functional characterization of mutants.

Members of the APOBEC3 (A3) cytidine deaminase family (*i.e.*, A3D, A3F, A3G, and A3H) can restrict human immunodeficiency virus (HIV) through lethal hypermutation of the HIV genome. The most potent HIV-1 restrictor is A3G (23, 24). Antiviral A3G activity requires that A3G binds to the HIV-1 RNA genome, packages into the virus particle, and has deaminase activity (25–27). HIV-1 counters this antiviral activity through HIV-1 accessory factor Vif, which binds to A3 proteins preventing their packaging and RNA binding. In addition, Vif recruits the Cullin-5 (Cul5) ring E3 ubiquitin ligase (CRL5) to target A3 proteins for poly-ubiquitylation and degradation by the 26S proteasome (28–31).

A3G has two domains: an N-terminal domain (NTD) that binds to single-stranded DNA (ssDNA) and RNA (ssRNA) and a catalytic C-terminal domain (CTD) with deaminase activity. While the NTD does not have any catalytic activity, it is required for several antiviral functions, including interaction with Vif (32), A3G dimerization (33, 34), processivity (35), subcellular distribution (36, 37), and packaging into the HIV-1 capsid (38). No atomic structure of the human A3G-Vif complex has yet been determined. Nevertheless, several key residues at the $A3G_{NTD}$-Vif interface have been identified by mutational, functional, evolutionary, and computational studies of various Vif and species homologs or mutant A3Gs (21, 39–43). However, these studies sometimes appear in conflict with each other; for example, studies have been inconclusive if A3G region 31 to 45 is necessary for binding the full-length and CTD of Vif (21, 44).

While the mechanism for Vif neutralization of A3 family members has been broadly characterized (26, 45, 46), the purification and structural characterization of Vif-CRL5 complexes has been hindered by their instability. Our previous studies that systematically mapped physical interactions between HIV and host proteins using affinity tag purification mass spectrometry (AP-MS) (47) revealed that Vif stably interacts with the cellular transcription factor core-binding factor β (CBFβ). Follow-up studies demonstrated that the CBFβ-Vif heterodimer binds to Elongin B/C complexes to act as the substrate receptor for the CRL5- E3 ligase and that CBFβ is required to stabilize the complex (30). Following this discovery, the structure of the stable Vif-CBFβ-EloB-EloC-$Cul5_{NTD}$ subcomplex was determined by X-ray crystallography (48) along with partial structures of the $A3G_{NTD}$, $A3G_{CTD}$, and Cul5 (supplemental Fig. S1). However, the structure of the full A3G substrate-bound complex, which consists of a full-length Cul5 bound to the E2 docking protein Rbx2, and the Vif-CBFβ-EloB/C (VCBC) adaptor complex, has been refractory to traditional structure determination techniques, likely because of its compositional and/or structural heterogeneity. Additional challenges that may have prevented the structural characterization of the A3G-Vif-CRL5-CBFβ complex include the low stability and solubility of the full-length A3G protein purified from *Escherichia coli* and the intrinsically transient nature of substrate–enzyme interactions.

In this work, we use the lysine-reactive MS-cleavable DSSO cross-linker (17) to obtain a comprehensive set of direct PPIs in the A3G-Vif-CRL5-CBFβ complex. Based on the cross-links, atomic structures of the components (supplemental Fig. S1), and previously published mutagenesis data, we then computed an integrative structure of the A3G-Vif-CRL5-CBFβ complex in solution. The structure was cross-validated based on the data used for modeling and previously published structural, biochemical, and functional data not used for modeling. This study has demonstrated the utility of DSSO-based XL-MS for mapping proximal lysine residues in structurally heterogeneous complexes, in turn enabling integrative structure determination of such complexes. The combined experimental and computational approach described here is generally applicable to other challenging host–pathogen protein complexes.

EXPERIMENTAL PROCEDURES

DSSO Cross-linking and Protein Digest

The complexes used in this study that contain either consensus Vif (Vif_{con}) or LAI Vif (Vif_{LAI}) (*i.e.*, A3G-Vif_{con}-CRL5-CBFβ, Vif_{con}-CBC-$Cul5_{NTD}$, Vif_{con}-CRL5-CBFβ, and A3G-Vif_{LAI}CBC) were purified as

previously described in (49–52). Reconstituted complexes were diluted to ~5 μM (different preparations at 1–1.2 mg/ml) in 20 mM HEPES pH 7.5, 300 mM NaCl, 10% Glycerol. Samples were reacted with increasing molar ratios of DSSO (supplemental Table S1) and cross-linking reactions carried out at 37 °C for 30 min at 1000 RPM on an Eppendorf Thermomixer C. All reactions were quenched with 100 mM NH_4HCO_3 or 100 mM Tris pH 8.0, then mixed with SDS-PAGE loading dye. Cross-linked samples were analyzed by SDS-PAGE and stained with MS safe blue stain (AcquaStain, Bulldog Bio). Bands corresponding to cross-linked proteins (as compared with non-cross-linked control samples) were excised and subjected to in gel digest with either trypsin or chymotrypsin (supplemental Fig. S2). Cross-linked peptides were analyzed by LC-MS^3 and identified through database searching as described below.

Analysis of Cross-linked Peptides by LC-MS^3

Dried peptide samples were dissolved in 2% FA, 3% ACN, and submitted for specialized LC-MS^3 analysis using an Easy-nLC 1000 (Thermo Fisher Scientific) coupled to an Orbitrap Elite Hybrid Mass Spectrometer with ETD (Thermo Fisher Scientific). Online peptide separation was performed with a 75 μm × 30 cm fused silica IntregraFrit capillary column (New Objective) packed in-house with 1.9-μm Reprosil-Pur C18 AQ reverse-phase resin (Dr Maisch-GmbH). Peptides were eluted at a flow rate of 300 nl/min using the following gradient of 5% B for 1 min, 5% to 35% B in 50 min, 35% to 95% B in 5 min, and 95% B for 4 min (mobile phase buffer A: 100% H_2O/0.1% FA; mobile phase buffer B: 100% ACN/0.1% FA). Each individually collected in-gel digest band sample was run in technical duplicate by two similar data-dependent acquisition methods for MS^3 analysis of cross-linked peptides (based on methods in (16)). For each method, a single acquisition cycle consisted of either 9 or 11 scan events as follows: 1) one full MS^1 scan in the orbitrap (350–1500 m/z, 120,0000 resolution, AGC target of 1×10^6, max injection time of 100 ms); 2) two data-dependent MS^2 scans in the orbitrap (15,000 resolution, AGC target of 5×10^4, max injection time of 500 ms) with normalized collision energy set at 22% on the top two precursor ions; and 3) either three or four MS^3 scans in the ion trap (ion count target 10^4, max injection time of 50 ms) with normalized collision energy set at 35% on the top three or four ions from each MS^2 scan. Precursors with charge state 4 and above were sampled for MS^2 and dynamically excluded for 20 s (tolerance of 10 ppm), with charge state and dynamic exclusion turned off for MS^3.

Identification of Cross-linked Peptides

Cross-link peptide identification was carried out in a semi-automated three-step process (supplemental Fig. S3). Briefly, MS^2 and MS^3 data were separately extracted from raw files using MSConvert (ProteoWizard) (53, 54). MS^3 data were searched on a locally installed version of ProteinProspector (v. 5.19.1) for peptide identification and DSSO-remnant modification localization. Exact peptide search and filtering criteria, including scores and E-values, for all trypsin and chymotrypsin data can be found in the *ProteinProspector Batch-Tag* and *searchCompare* setup files, which are provided as xml or txt supplemental Files (bt_trypsin.txt, bt_chymotrypsin.txt, bt_trypsin.xml, bt_chymotrypsin.xml, sc.xml, and sc.txt). MS^2 and MS^3 data and peptide identifications are integrated and cross-linked peptides (*e.g.*, dead-end or mono-linked, loop-linked, and interlinked) were identified by XLTools, a revised version of XL-Discoverer (55, 56). Cross-linked peptide data were then summarized with ambiguous assignments deconvoluted and unique and redundant peptides quantified and scored using in house scripts (available at https://github.com/integrativemodeling/A3G-CRL5-Vif-CBFb). For detailed methods and calculations, see Supplementary Information. All of the proteomics datasets from each step, including raw files, MS^2 and MS^3 peak files, and MS^3 search files, have been deposited to the ProteomeXchange Consortium *via* the PRIDE (57) partner repository with the dataset identifier PXD025391. Annotated spectra for all interlinked, loop-linked, dead-end, and single peptides can be found on the MSViewer application through ProteinProspector (58) (https://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) with the search key 9tjmaqhszr.

Integrative Structure Modeling

We applied an integrative structural modeling approach to characterize the structure of the A3G-Vif-CRL5-CBFβ complex in solution, based on the atomic structures of the components, 132 DSSO cross-links, and previously published mutagenesis data. Integrative structure determination proceeded through the standard four stages (19, 20, 59): 1) gathering data, 2) representing subunits and translating data into spatial restraints, 3) sampling of configurations to produce an ensemble of structures that satisfies the restraints, and 4) analyzing and validating the ensemble structures and data. The integrative structure modeling protocol (*i.e.*, stages 2, 3, and 4) was scripted using the Python Modeling Interface (PMI) package, a library for modeling macromolecular complexes based on our open-source Integrative Modeling Platform (IMP) package (19). Detailed methods can be found in Supplementary Information.

Validation by Randomly Removing Data and Resampling

The resulting ensemble of models was validated by recomputing it with subsets of the data (*i.e.*, "jackknifing") (60). We performed sampling runs without the cross-links or mutagenesis data and with only random fractions of the cross-links.

Data Visualization

The sequence coverage and distributed count of all dead-end, loop-linked, and interlinked (intra- and intersubunit) cross-linked peptide data were visualized using a newly developed tool, XLs Summary Viewer implemented using the Matplotlib Python Package (61) and available in https://github.com/integrativemodeling/A3G-CRL5-Vif-CBFb. All protein structures and integrative models were visualized using Chimera (https://www.rbvi.ucsf.edu/chimera/) (62). MS^3 spectral data were visualized using the interactive Peptide Spectral Annotator (IPSA-http://www.interactivepeptidespectralannotator.com/BulkDataUpload.html) (63). Interlinked residue data were visualized as a circos linkage map using CX-Circos (v 0.11.1, http://www.cx-circos.net/), from data files automatically generated during step 3 of the cross-link peptide identification pipeline (supplemental Table S2).

Experimental Design and Statistical Rationale

Four HIV-1 Vif-containing complexes were expressed, purified, and reconstituted for *in vitro* cross-linking with disuccinimidyl sulfoxide (DSSO, (17)) (Fig. 1). These four complexes include the full-length A3G-Vif$_{con}$-CRL5-CBFβ complex (*i.e.*, A3G, Vif$_{con}$, CBFβ, EloB, EloC, Cul5, and Rbx2), and Vif-containing subcomplexes Vif$_{con}$-CRL5-CBFβ, Vif$_{con}$-CBC-Cul5$_{NTD}$, and A3G-V$_{LAI}$CBC (the number of replicates for each complex and the full experimental details are provided in supplemental Table S1). Each sample was prepared using multiple cross-linking concentrations to obtain a comprehensive set of cross-linked peptides. Each sample cross-linked at different concentrations was separated by SDS-PAGE and cross-linked bands excised for in-gel digest. Samples were digested by trypsin or chymotrypsin in gel, and the digest samples were analyzed by two similar LC-MS^3 methods to identify dead-end (mono-linked), loop-linked, and interlinked peptides (supplemental Table S1). The full-length A3G-Vif$_{con}$-CRL5-CBFβ complex was prepped five times, three times for trypsin

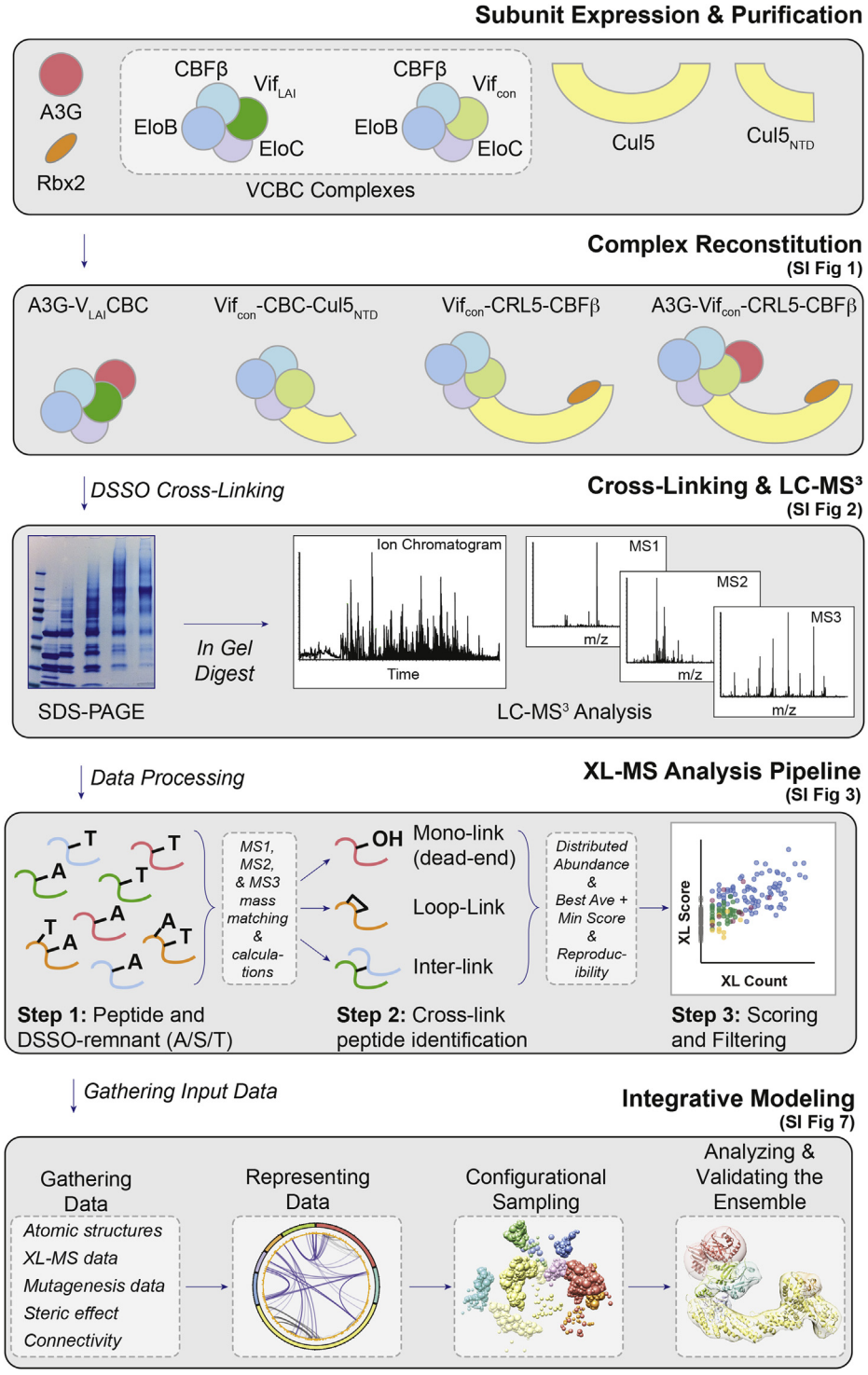


FIG. 1. **Overview of the pipeline for integrative modeling of structurally heterogeneous complexes based on cross-linking mass spectrometry (XL-MS).** First, all protein subunits are expressed or coexpressed and purified. Second, the full protein complex and selected stable subcomplexes are reconstituted *in vitro*. For more details on the subunit components and functions, see supplemental Fig. S1. Third, each protein complex is cross-linked using MS-cleavable DSSO cross-linker. Cross-linked protein complexes are then separated by SDS-PAGE and stained gel pieces are excised for in-gel enzymatic digestion (supplemental Fig. S2). The resulting peptide mixtures are separated and analyzed *via* liquid chromatography–tandem mass spectrometric (LC-MS) analysis. Fourth, database searching of the MS data identifies DSSO-modified peptides and their linkage sites. Identified interlinked, loop-linked, and mono-linked (dead-end) modified peptides are quantified, scored, and filtered by our XL-MS analysis protocol (supplemental Fig. S3). Lastly, the cross-link dataset and other available data are used for integrative modeling (supplemental Fig. S7).

digest (1-A to 3-A in supplemental Table S1) and twice for chymotrypsin digest (1-B and 2-B in supplemental Table S1) to extend cross-linked peptide coverage of the individual components in the complex. The Vif$_{con}$-CRL5-CBFβ and Vif$_{con}$-CBC-Cul5$_{NTD}$ complexes were prepared three and two times for trypsin digest, respectively. The A3G-V$_{LAI}$CBC was prepared only once for trypsin digest. The integrative structures were validated using a series of tests to determine precision and how well they agree with the input information used and not used for modeling (64, 65).

RESULTS

Pipeline for Integrative Modeling of Structurally Heterogeneous Host–Pathogen Complexes Based on XL-MS

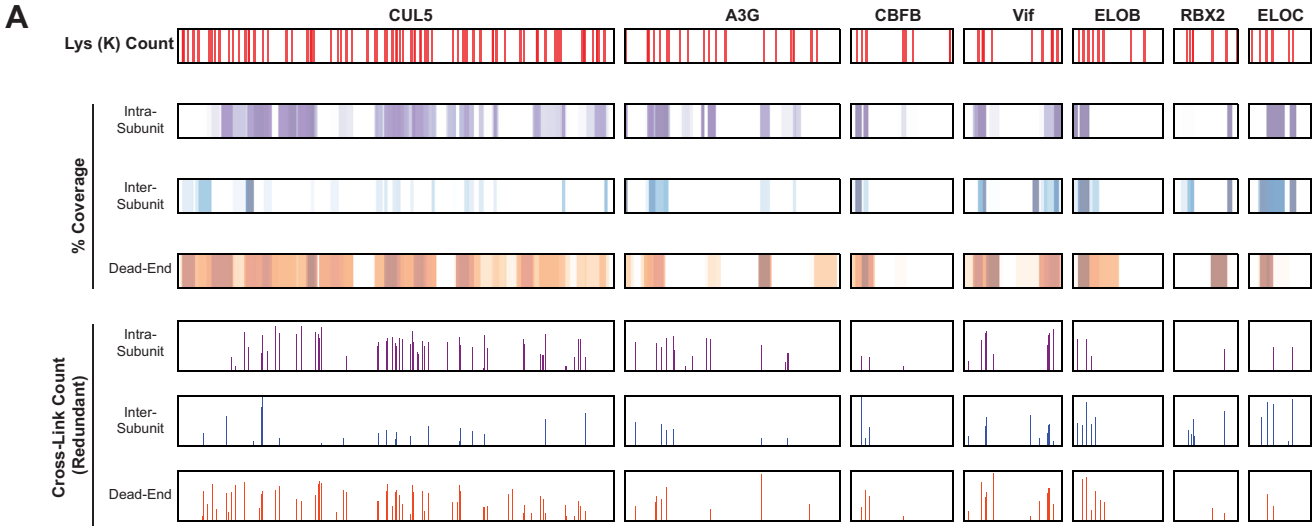
To characterize the architecture of host–pathogen complexes, we present a pipeline that streamlines integrative structure modeling based on chemical cross-link data and predicted residue–protein contacts from mutagenesis studies (Fig. 1). The general workflow includes expression, purification, and reconstitution of protein complexes, and XL-MS, and integrative structure modeling. To improve the sequence coverage of each subunit and the number of cross-links detected by MS, we included *in vitro* stable subcomplexes in addition to full-length protein holo-complexes. Each reconstituted protein complex is cross-linked using the MS-cleavable, amine reactive chemical cross-linker DSSO (17) that yields a complex mixture of: 1) unmodified peptides; 2) dead-end (or mono-linked) peptides; 3) loop-linked peptides, where two residues on the same intact peptide are linked; and 4) interlinked peptides, where two residues on two individual peptides (referred to as α and β) are linked (supplemental Fig. S4, *A* and *B*). Using specialized multistage liquid chromatography–mass spectrometry (LC-MS^n) methods, these four types of peptides are identified and measured (see supplemental Fig. S4, *C–F* for representative interlink peptide).

Cross-linked peptides that are identified more frequently (*i.e.*, they have a higher redundant count) and where both peptides (*i.e.*, α and β) individually have high scores (*i.e.*, the average and minimum of these two scores are both high) (SI Methods, (66, 67)), represent higher confidence linkages than those with low count and low scores. To quantify this confidence for use in integrative modeling, we define a composite score based on the frequency with which the linkage was identified (redundant count) as well as the highest average and minimum scores for the best representative peptides for each unique linkage. For ambiguous linkages (*i.e.*, cross-linked peptides with missed cleavages and sequentially close lysine residues), the counts were uniquely distributed based on the XL-remnant assignments and proportionally added to the unique residue–residue linkage events count (SI Methods). This deconvolution is inspired by the definition of the distributed normalized abundance factor (dNSAF) (68) that similarly distributed ambiguously assigned peptide spectral counts (based on homologous sequences within a database).

In addition to the cross-links data, we explored how integrative structure determination can benefit from spatial restraints derived from mutagenesis data indicative of interface residues. To this end, we converted the mutagenesis data into an upper bound on the distance between the residues identified to be required for binding and the closest residues in the predicted bound protein (SI Methods). We named this spatial restraint the *residue–protein proximity restraint*. An advantage of this approach is that it does not make assumptions about the orientation of the proteins or the binding interface on the predicted bound protein. Similar restraints have been implemented by other interactive modeling software to incorporate data from mutagenesis into docking protocols (69, 70).

XL-MS Analysis of A3G-Vif-CRL5-CBFβ

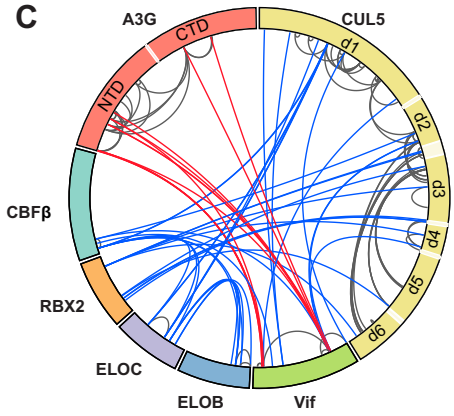
The full-length A3G-Vif$_{con}$-CRL5-CBFβ complex (*i.e.*, A3G, Vif$_{con}$, CBFβ, EloB, EloC, Cul5, and Rbx2), as well as three different Vif-containing subcomplexes (*i.e.*, Vif$_{con}$-CRL5-CBFβ, Vif$_{con}$-CBC-Cul5$_{NTD}$, and A3G-V$_{LAI}$CBC), were cross-linked with DSSO, digested, and analyzed by XL-MS^3 for peptide and cross-link identification. The subcomplexes allowed us to compare contributions of each protein, look at two different Vif protein variants (*i.e.*, LAI and consensus), and helped to increase our confidence in A3G and Vif-CBC cross-links. We obtained a good overlap of detected cross-links between the A3G-V$_{LAI}$CBC and A3G-Vif$_{con}$-CBC subcomplexes (supplemental Fig. S5). Consequently, the data of all subcomplexes were combined. Reproducible (*i.e.*, uniquely identified in at least in two separate MS runs either from a biological or technical replicate) peptides were included for further cross-links processing and integrative modeling. In total, we identified 6959 dead-end, 6300 interlinked, and 389 loop-linked redundant (nonunique) peptides from all seven components of the A3G-Vif-CRL5-CBFβ complex (supplemental Tables S2–S6). The estimation of the false discovery rate (FDR) was 2.1%, as determined by the percentage of cross-linked peptides for which at least one partner is from the decoy database (1.4%) or from a non-A3G-Vif-CRL5-CBFβ proteins (0.7%). Interlinked peptides had a slightly higher FDR (3%) as compared with dead-end (1%) or loop-linked (1.3%) peptides. Through pooling the data collected from A3G-Vif-CRL5-CBFβ and the other three Vif-containing subcomplexes, we achieved 72% coverage of the protein sequences and 83% coverage of the reactive residues in the full A3G-Vif-CRL5-CBFβ complex (Fig. 2*A*). This coverage includes dead-end modifications and interlinks identified for all seven proteins, indicating good coverage for each individual subunit within the complex. Different subunits demonstrate different proportions of dead-end and intrasubunit or intersubunit linkages (Fig. 2*B*). For instance, A3G shows a higher proportion of dead-end and intrasubunit linkages than intersubunit linkages, indicating that A3G might be involved in less stable PPI interfaces as compared with CBFβ,



B

Table 1. Redundant Count of DSSO Cross-Linked Peptides

Subunit	Dead-End (Mono-Link)	Intra-Subunit (Loop-Link)	Intra-Subunit (Inter-Link)	Inter-Subunit (Inter-Link)
A3G	993	5	327	97
CBFβ	123	0	9	1047
CUL5	3315	275	2299	1723
ELOB	951	0	97	593
ELOC	50	0	28	1623
RBX2	9	0	11	190
VIF	1435	104	591	203



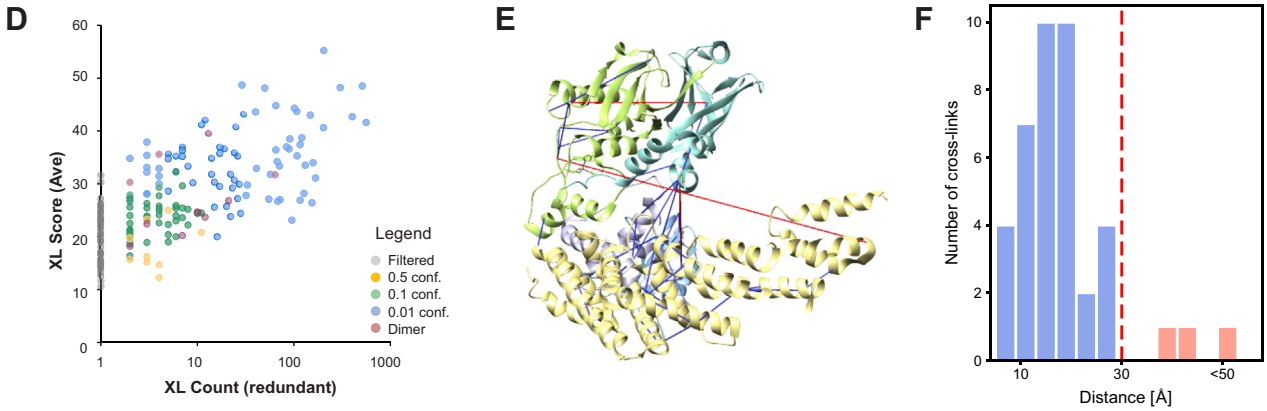


FIG. 2. **DSSO XL-MS analysis of the A3G-Vif-CRL5-CBFβ complex and subcomplexes.** *A*, overview of cross-linked peptide coverage for all seven proteins in A3G-Vif-CRL5-CBFβ. The position of each lysine (Lys Pos) is shown on the *top panel* with *red lines*. The sequence coverage (% Cov) and redundant cross-link count (XL Count) of each protein are shown below. Residue positions in intrasubunit (*purple*), intersubunit (*blue*), and dead-end (*orange*) linkages are also shown below. *B*, table summarizing the number of redundant cross-links identified for each subunit. *C*, CX-Circos cross-links map of all A3G-Vif-CRL5-CBFβ unique cross-links. Intra- and intersubunit cross-links are represented by *gray* and *blue edges*, respectively. A3G to Vif cross-links are colored *red*. The A3G NTD and CTD rigid bodies used for integrative modeling are labeled on A3G as *darker pink* with the flexible linker being *lighter pink*. The Cul5 domains used as rigid bodies in the dynamic model are labeled on Cul5 as *darker yellow* and the flexible linkers are in *light yellow*. *D*, scatter plot showing the relation between redundant count and the cross-link average alpha and beta peptides identification scores. A high (*blue*), medium (*green*), or low (*orange*) composite confidence score was assigned to each cross-link. Cross-links identified only once (redundant count = 1) or cross-links

EloC, or Rbx2, which demonstrate higher percentages of intersubunit linkages (Fig. 2*B*).

Based on the identified cross-linked peptides (supplemental Fig. S5), unique residue-to-residue linkages were determined. To ensure the validity of subsequent analyses, we use high scoring and reproducible residue-to-residue linkages that are observed across biological replicate experiments (SI Methods). A total of 132 reproducible and high confidence cross-links for A3G-Vif-CRL5-CBFβ were classified into three confidence categories and used for subsequent analysis and modeling (supplemental Table S2 and Fig. 2, *C* and *D*).

Of the 132 high-confidence A3G-Vif-CRL5-CBFβ cross-links identified, only 30% (40) could be mapped to the X-ray structure of the Vif-CBC-Cul5$_{NTD}$ subcomplex, with the remaining cross-links mapping to either missing segments or missing subunits (Fig. 2*E*). This mapping revealed that 37 cross-links (93%) are satisfied by the X-ray structure, where satisfied cross-links are defined as having a distance between the Cα of the cross-linked residues less than 30 Å (Fig. 2*F*) (71). This high cross-link satisfaction suggests that the X-ray crystal structure of the Vif-CBC-Cul5$_{NTD}$ subcomplex is similar to the full-length A3G-Vif-CRL5-CBFβ structure in solution. Furthermore, we observed a large number of cross-links between regions known to be at or near the binding interface between Vif and Cul5 and CBFβ and Cul5 (supplemental Fig. S6). The cross-links that are violated may be explained by the conformational heterogeneity of the A3G-Vif-CRL5-CBFβ complex, differences between solution and crystallographic structures, or uncertainty in mass spectrometry identifications. Similarly, we mapped the cross-links to the structure of the Vif-CBFβ-Cul5-Rbx2-EloB-EloC subcomplex (Vif-CRL5-CBFβ subcomplex) for which the Cul5 CTD and Rbx2 subunit were modeled using comparative modeling (72, 73) (SI Results). In this case, 87 (66%) of the 132 unique cross-links mapped to the Vif-CRL5-CBFβ modeled subcomplex and 80% of the cross-links are satisfied (supplemental Fig. S7). The unsatisfied cross-links mostly span residues between the Cul5 NTD and Vif-CBC complex or between Cul5 and Rbx2, indicating that the comparative model of the full-length Cul5 does not fully capture its solution structure or its structural heterogeneity. In addition, 34% of the identified cross-links cannot be mapped to either the Vif-CBC-Cul5$_{NTD}$ X-ray structure or the Vif-CRL5-CBFβ comparative structure, due to segments or components missing from the structure. Thus, we applied integrative structure modeling to determine the full structure and structural dynamics of the A3G-Vif-CRL5-CBFβ complex.

The Integrative Structure of A3G-Vif-CRL5-CBFβ

To determine the A3G-Vif-CRL5-CBFβ solution structure, we performed integrative structure modeling using the previously described four-stage workflow (SI Methods and supplemental Fig. S8) (7, 19, 20, 74). The input information includes our scored and filtered 132 DSSO cross-links, the X-ray structure of the Vif-CBC-Cul5$_{NTD}$ subcomplex (48), comparative models of the A3G NTD and CTD, and full-length Cul5 and Rbx2 (SI Methods), and predicted residue–protein contacts from mutagenesis studies (21, 22, 41, 75).

Preliminary integrative modeling represented the Vif-CRL5-CBFβ subcomplex by a single rigid body defined by atomic structure of the Vif-CBC-Cul5$_{NTD}$ subcomplex (48) and comparative models of the full-length Cul5 and Rbx2. Each domain of A3G was represented as a separate rigid body. The linker between the A3G domains (residues 195–200) and regions missing in the X-ray structure were represented by flexible strings of beads (*i.e.*, one residue-per-bead). This representation and output models are hereafter referred to as the rigid representation and model. Models obtained using this representation did not satisfy the spatial restraints implied by the data within the uncertainty of the data (below, SI Methods, supplemental Table S7). Therefore, to characterize the flexibility of the complex, we performed integrative structure modeling of the A3G-Vif-CRL5-CBFβ complex with a modified representation that allowed for relaxation in the configuration of the Vif-CBC subcomplex subunits and for alternative conformations of Cul5 and A3G subunits. Specifically, we represented each protein of the Vif-CBC subcomplex and Rbx2 as independent rigid bodies. Cul5 was represented as six rigid bodies connected by flexible linkers (4–23 residues each) representing the loop between the: (i) second and third repeat of the NTD, (ii) elongated NTD and globular CTD, (iii) 4-helix bundle, (iv) α/β, (v) WH-A, and (vi) WH-B domains (SI Methods, supplemental Fig. S9*A*) (76). The A3G NTD and CTD remain represented as rigid bodies (supplemental Table S8). This representation and output models are hereafter referred to as the flexible representation and model.

With these representations in hand, we next translated the input information into spatial restraints as follows. First, the 132 DSSO cross-links were used to construct a Bayesian term that restrained the distances spanned by the cross-linked residues (SI Methods). Second, we converted the mutagenesis data into residue–protein proximity restraints. Two such restraints were defined corresponding to well-established genetically confirmed regions required for A3G binding and degradation, namely A3G residues 126 to 132 and Vif residues 40 to 45 (21, 22, 41, 75). Third, to use the crystal structure of the Vif-CBC-Cul5$_{NTD}$ subcomplex and comparative model of Cul5/Rbx2 as templates, we imposed *structural equivalence distance restraints* designed to restrain the model to resemble

corresponding to homodimers were filtered out (*gray*). The composite confidence scores were used for integrative structure modeling. *E*, satisfied (*blue*) and violated (*red*) cross-links are mapped onto the X-ray structure of the Vif-CBC-CUL5$_{NTD}$ subcomplex (PDB code 4N9F), using a 30 Å Cα-Cα distance cutoff. *F*, histogram of the distances between cross-linked residues in the X-ray structure. The *vertical dashed red line* represents the 30 Å cutoff used to define satisfied XLs (Cα-Cα distance < 30 Å).

the templates as much as possible (77). These restraints were applied between pairs of residues closer than 7.0 Å across an interface between two rigid bodies only in the flexible representation. Lastly, we applied sequence connectivity and excluded volume restraints to all components (SI Methods).

Structural models of the A3G-Vif-CLR5-CBFβ complex were obtained by exhaustive Monte Carlo sampling, starting with random initial configurations. The clustering of the ensembles identified a single cluster of models with 91% and 98% of the individual structures for the rigid and flexible models, respectively (Fig. 3, *A* and *B*). The model precision, which is defined as the average RMSD between all solutions in the ensemble, is 8.4 and 19.9 Å for the rigid (supplemental Fig. S10) and flexible model, respectively; this variability arises both from the actual structural heterogeneity and relative lack of input information for structure modeling. In the rigid model, 89% of the cross-links are satisfied by at least one model in the structural ensemble, including all of the cross-links attributed to A3G (Fig. 3*C* and supplemental Fig. S9*B*). Unsatisfied cross-links span mostly residues between the Cul5 CTD and the Vif-CBC subcomplex, indicating that the rigid representation of the system is not adequate to capture the full range of conformations in solution. In contrast, the flexible model satisfies 99% of the cross-links (Fig. 3*C* and supplemental Fig. S9*C*). Similarly, residues included in the residue–protein proximity restraint are within a threshold distance of the binding target in the rigid and flexible models (Fig. 3*D*). This uncertainty in the output model primarily reflects the structural heterogeneity of the sample, the lack of information, or a combination of both (below). These ensembles allow us to identify the A3G binding conformation and to quantify the A3G and Cul5 structural heterogeneity within the model ensembles. For most aspects the conclusions do not depend on the representation (*i.e.*, rigid or flexible) used. For

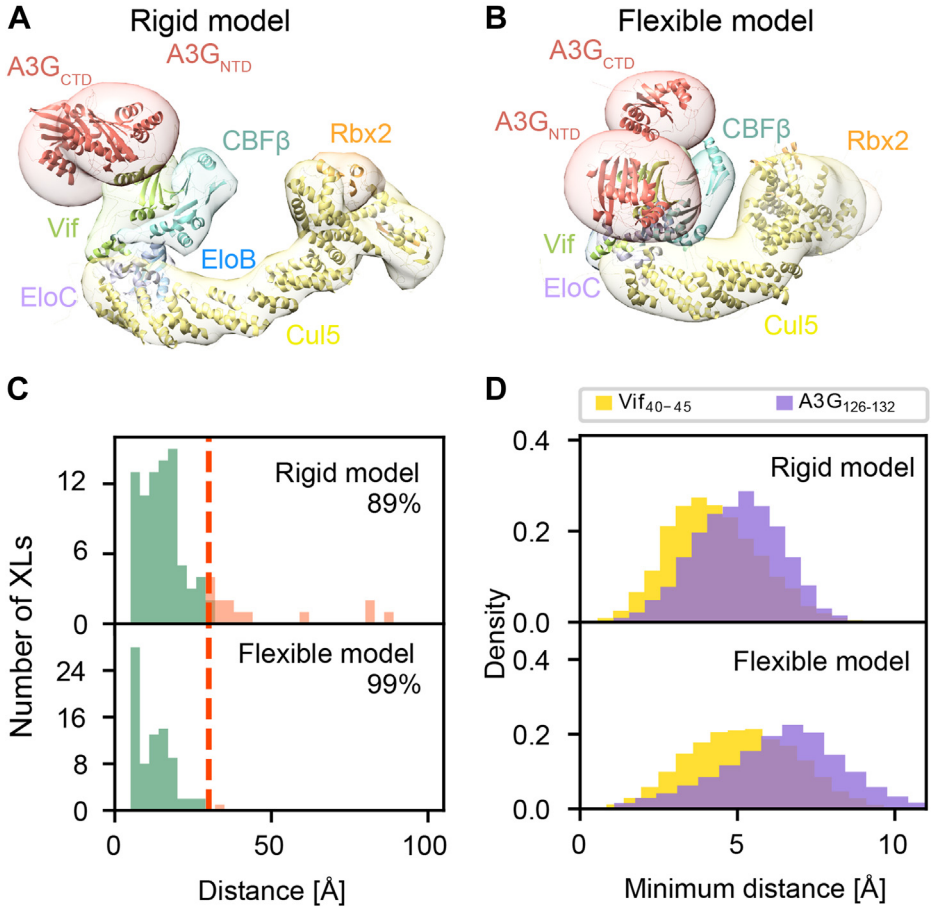


FIG. 3. **Integrative structure of the A3G-Vif-CRL5-CBFβ complex.** *A*, structure of the A3G-Vif-CRL5-CBFβ complex using a rigid representation. The localization probability density of the ensemble of structures is shown with a representative (centroid) structure from the ensemble embedded within it. A localization density map for a set of models is defined as the probability of observing a model component at any point in space. *B*, structure of the A3G-Vif-CRL5-CBFβ complex using a flexible representation. *C*, histogram showing the distribution of the cross-linked Cα–Cα distances in the A3G-Vif-CRL5-CBFβ integrative structures (*top*: rigid representation and *bottom*: flexible representation). Cross-links with Cα–Cα distances that fall within the distance threshold of 30 A in at least one structure in the ensemble are classified as satisfied. *D*, histograms showing the minimum distances between the selected segments and the target protein surfaces for each of the protein-surface proximity restraint.

example, we obtained the same A3G-Vif interface with the rigid and flexible representations (supplemental Fig. S11).

To validate the model and to determine if the data types used are consistent with each other, we computed the model using subsets of the data. For example, to determine the effect of the two residue–protein proximity restraints on both the model and its precision, we recomputed the model of the A3G-Vif-CRL5-CBFβ complex excluding these restraints using the rigid representation (supplemental Fig. S10). As expected, the resulting model is less precise (model precision of 13.7 Å), although it has a similar cross-link satisfaction (89%) than the model computed with the residue–protein proximity restraints. The model computed without the residue–protein proximity restraints has the same architecture as the original model; in particular, the A3G-Vif interface is conserved within the precision of the model (supplemental Fig. S11*B*). As another validation, we recomputed the structure including random subsets of cross-links (*i.e.*, "jackknifing") (60) (supplemental Fig. S11). Removing a small fraction (20%) of the cross-links had no effect on the model precision, but removing larger fractions resulted in less precise models. The obtained A3G-Vif interfaces are similar to the original interface (*i.e.*, rigid model), suggesting that the cross-link data are accurate, and the models are not a result of overfitting. This result is expected when the uncertainty in the output model reflects primarily structural heterogeneity of the sample, not the lack of data. Finally, we recomputed the model without discriminating among the cross-links based on their composite confidence scores (supplemental Fig. S11). This is equivalent to just filtering the cross-links based on reproducibility and a threshold score. The resulting model has two clusters (populations of 50% and 25%), is less precise (model precisions 10.5 and 9.1 Å) than the original model, although each cluster satisfies the cross-links (89%) as well as the original model. The A3G-Vif interfaces in these clusters are similar to the original interface, indicating that cross-link composite confidence scores improves the precision of the resulting models. In conclusion, increased precision of the model resulting from using a larger number of restraints also increased our confidence in the model and the data alike (subject to sufficient sampling, SI Methods, supplemental Tables S7 and S8).

The Structure Reveals That A3G Binds to Vif Mostly Through Its NTD

Normalized contact frequencies, defined by how often in the ensemble any pair of residues contact each other in the ensemble, identified a single A3G-Vif interface (Fig. 4, *A* and *C*). Mapping of the normalized contact frequencies to the protein surfaces revealed that the A3G NTD and CTD domains interact mostly through the 241 to 258 loop (Fig. 4*B* and supplemental Fig. S12*A*). The A3G-Vif interface contains the residues identified in mutagenesis-based studies, including the residues restrained by residue–protein proximity restraint (*i.e.*, A3G 126–132 and Vif 40–45) as well as other regions, such as A3G loops α1-β1, β2-α2, and β4-α4 (75) and Vif residues 22 to 26 (32) (Fig. 4*B*). In addition to these previously described A3G-Vif contacts, we also identified A3G regions 30

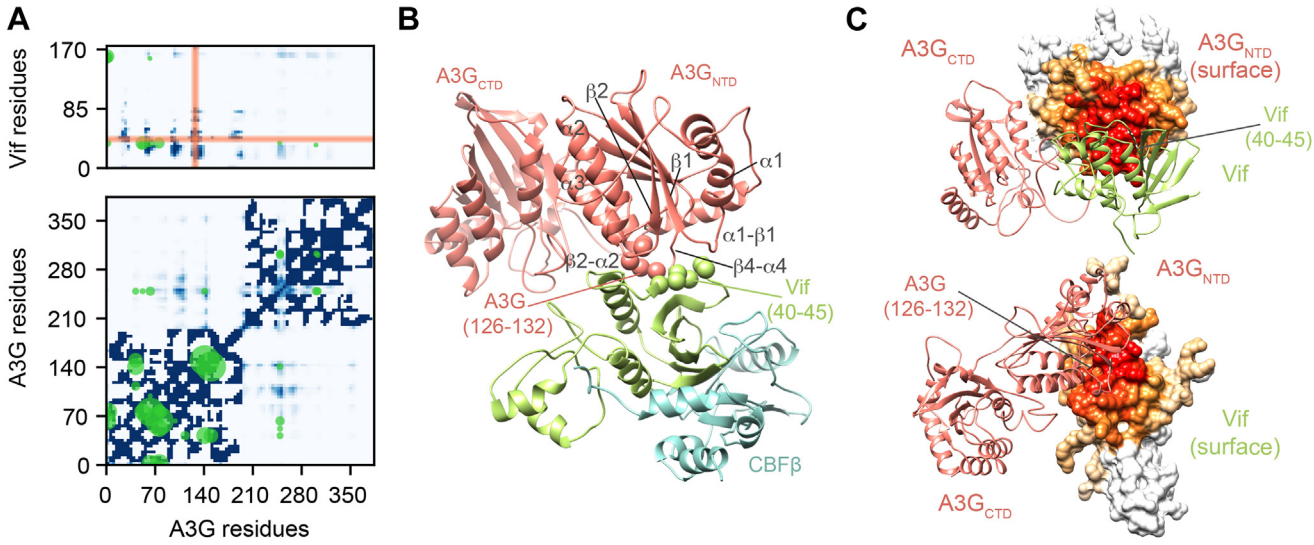


FIG. 4. **A3G binding to the CRL5-Vif-CBFβ complex.** *A*, contact maps computed for A3G and the A3G-Vif interface. The *blue bins* indicate pairs of proximal beads representing the model, with the intensity of blue proportional to the fraction of models in the cluster whose Cα–Cα distance is closer than the cutoff of 12 Å. The *green circles* correspond to the cross-links satisfied by at least one model in a cluster, with the size of the circle proportional to the spectral count. The *orange vertical* and *horizontal regions* correspond to the segments selected for the residue–protein proximity restraint. *B*, detail of the A3G -Vif binding interface. Segments A3G 126 to 132 and Vif 40 to 45 are shown as spheres. *C*, A3G-Vif binding interface. Surface representation of Vif (*top*) and A3G (*bottom*) showing the binding interface with the intensity of red proportional to the fraction of models in the cluster whose Cα–Cα distance is closer than the cutoff of 12 Å.

to 35, 55 to 66, 92 to 101, and 188 to 195 as part of the A3G-Vif interface (supplemental Table S9).

The Integrative Structure of A3G-Vif-CRL5-CBFβ Reveals That the A3G CTD and Cul5 Subunits Are Structurally Dynamic

The A3G CTD is not localized precisely in our model. We quantified the configurational and conformational heterogeneity of the A3G NTD and CTD domains using two structural metrics. First, we computed the RMSD of each domain with respect to the Vif-CBC subcomplex and with respect to each other (Fig. 5*A*). After superimposing the Cα coordinates of the Vif-CBC complex, the average RMSD of the A3G NTD and CTD domains is 4.5 ± 1.8 Å and 7.0 ± 2.0 Å, respectively. Similarly, the RMSD of the CTD domain with respect to the NTD domain is 12.8 ± 3.8 Å. Second, we computed the solvent-accessible surface area (SASA) for the A3G NTD-Vif, A3G CTD-Vif, and A3G NTD-CTD interfaces (Fig. 5*B*). Together, computed RMSDs and SASAs indicate that the A3G CTD (i) adopts a range of conformations, (ii) does not have a significant binding interface with Vif, and (iii) is only loosely associated with the A3G NTD. The A3G NTD-CTD interaction occurs mostly through the 241 to 258 loop (supplemental Fig. S12*A*). This conformation has previously been described as the dumbbell form, for free (78) and ssDNA-bound A3G (79).

In the flexible model ensemble, Cul5 adopts a closed conformation, bringing the Cul5 CTD closer to the Vif-CBC complex (Fig. 3*B*). Importantly, the proteins in the Vif-CBC-Cul5$_{NTD}$ subcomplex adopt configurations that are similar to the one in the X-ray structure. Configurational changes include the rotation of EloB, EloC, and Cul5 with respect to Vif (supplemental Fig. S12*B*). The average RMSD between the X-ray structure and the models of Vif-CBC-Cul5$_{NTD}$ in the ensemble is 12.1 Å (supplemental Fig. S12*C*).

To indicate the most flexible parts of the structure, we assessed the uncertainty of the position and orientation of each rigid body in the flexible model ensemble. To this end, all models were superimposed on each rigid body in turn, followed by computing the average RMSD for each of the other rigid bodies (Fig. 5*C*). The model ensemble indicates large variability in the positions and orientations of some of the Cul5 domains with respect to the Vif-CBC complex. In particular, the Cul5 1 to 302 region presents a lower RMSD variability with respect to the Vif-CBC complex, while a large variability is observed for the globular Cul5 CTD. This observation is consistent with the structural heterogeneity of this region indicated by the relative lack of electron density from crystallography and molecular dynamics studies that indicate the Cullins are flexible proteins with conserved hinges in the NTD (80, 81). Furthermore, it has been shown that Neddylation of Cul5 leads to a conformational change in the Cul5 CTD (82, 83). No Nedd8 was present in our sample, but we hypothesize that a similar range of conformations is available for the sample in solution.

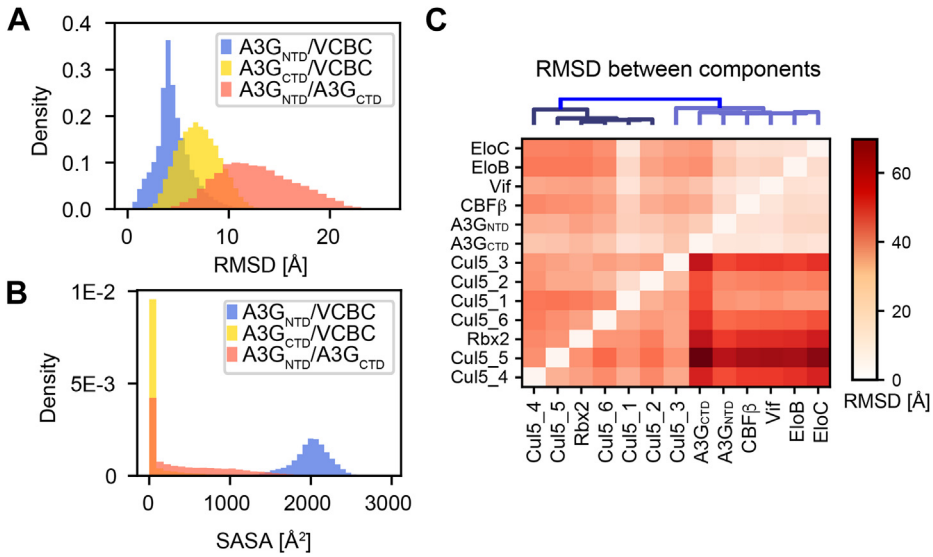


FIG. 5. **Structural heterogeneity of the A3G-Vif-CRL5-CBFβ complex.** *A*, distribution of RMSD between the A3G domains and Vif-CBC subcomplex after superimposing the Cα coordinates of Vif-CBC computed for all structures in the ensemble. *B*, distribution of SASA for the A3G NTD-Vif, A3G CTD-Vif, and A3G NTD-CTD interfaces computed for all structures in the ensemble. *C*, RMSD between rigid bodies in the model ensemble. The *vertical axis* corresponds to the rigid body used as reference for superimposition and the *horizontal axis* are the rigid bodies for which the average RMSD was computed. Cul5 segments numbered 1 to 6 correspond to the rigid bodies between the flexible linkers (supplemental Table S8).

The A3G Structure and A3G-Vif Interface Allow Us to Rationalize Structural, Biochemical, and Functional Data

We also quantified the degree to which the model satisfies relevant information not used to compute it, including data from structural, biochemical, and functional studies. First, we compared our integrative structure of A3G bound to the Vif-CRL5-CBFβ complex to the recently published X-ray structure of the full-length double-domain rhesus macaque A3G (rA3G) (84). The rA3G structure revealed that the two domains interact through a flexible linker (residue 194–198 in rhesus macaque and 195–199 in humans). However, two rA3G constructs differing in several point mutations have a different packing orientation between the two rA3G domains. To compare the X-ray structures to the integrative modeling ensemble, we computed the distribution of the Cα root-mean-square deviation (RMSD) between the X-ray structures and each of the models in the ensemble (Fig. 6*A*). The mean/minimum Cα RMSDs are 9.4/3.6 and 10.0/3.5 Å for the 6P40 and 6P3X structures, respectively (84). The Vif-CRL5-CBFβ bound A3G structure superimposes well on the rA3G structure and shares the same domain interface, mediated by the 243 to 258 flexible linker (Fig. 6*B*). The RMSD distributions show that the integrative structure ensemble includes the X-ray structures within the precision of the model and captures the overall conformational heterogeneity of the domain orientations. Thus, the integrative structure is validated by the X-ray structure, even though the X-ray structure covers only one out of seven protein subunits.

Second, we examined how the A3G and Vif regions that have been indicated as having functional effect by mutagenesis studies map to the A3G-Vif interface on the integrative model. The A3G-Vif interaction has been studied extensively, mainly through monitoring the mutation effect on A3G degradation in cells, A3G packaging into viral particles, A3G-Vif binding stability, and HIV restriction. To this end, we prepared a comprehensive overview of the A3G and Vif mutations and their functional effects (supplemental Tables S9 and S10, supplemental Fig. S13) and map these regions to the A3G-Vif interface on our integrative model (Fig. 6, *C* and *D*). Specifically, we mapped A3G and Vif mutations shown to: 1) disrupt

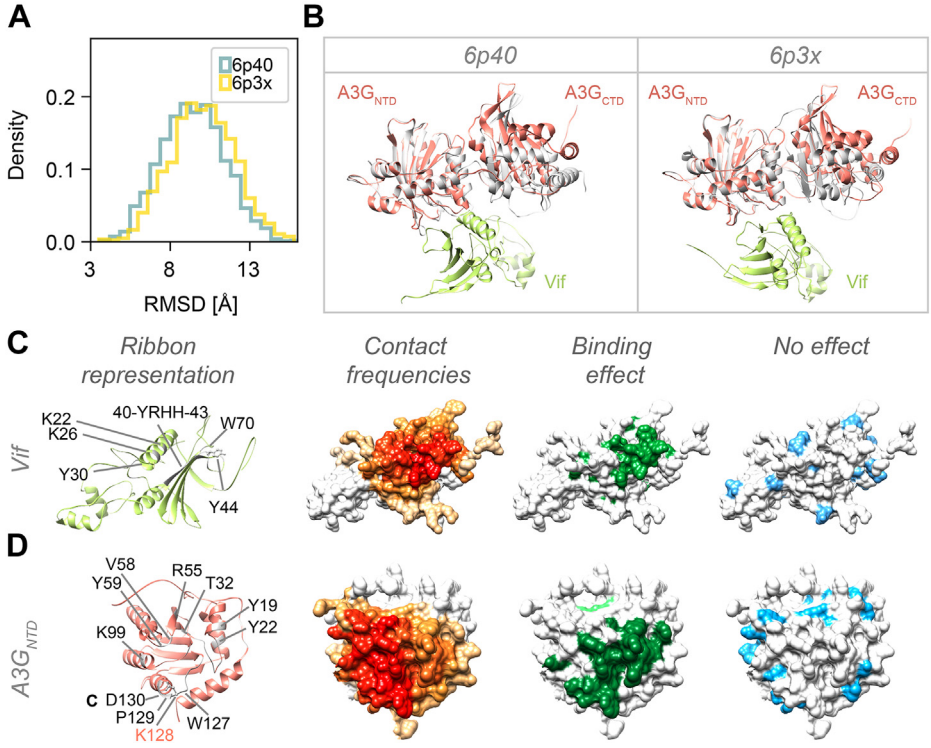


FIG. 6. **Rationalization of the A3G structure and A3G-Vif Interface by structural, biochemical, and functional data.** *A*, comparison of the A3G-bound to Vif-CRL5-CBFβ integrative ensemble to the rA3G X-ray structures (PDB codes 6p40 and 6p3x) (84). Histogram showing the Cα RMSD between A3G in all the model ensemble and the two rA3G X-ray structures. *B*, superimposition of the A3G-bound integrative (*colored*) and rA3G (*gray*) X-ray structures. The superimposition is based on the A3G NTD to show the rotation of the A3G CTD. Vif (*green*) is shown for reference. *C*, residue on Vif involved in A3G binding. From *left* to *right*: *ribbon* representation of Vif showing key residue that is known to be required for A3G-Vif binding; surface representation of Vif showing the normalized contact frequencies for the Vif-A3G NTD interface. The intensity of red is proportional to the normalized contact frequency of models in the integrative ensemble whose Cα–Cα distance is closer than the contact cutoff of 12 Å; surface representation of Vif showing in *green* the residues involved in A3G binding identified in mutagenesis studies; and surface representation of Vif showing in *cyan* the residues that had not effect on A3G binding in mutagenesis studies. *D*, residue on A3G NTD involved in Vif binding. Same as panel *C*.

A3G-Vif binding (21, 22, 44, 75), 2) protect A3G from degradation in cells, 3) restore packaging into the virions or virus-like particles (40, 85–87), 4) restore A3G restriction activity by hypermutation of the HIV-1 genome (40, 88). Additionally, we mapped mutations that have no demonstrated functional effect. Residues or regions where studies have seemingly contradictory data (*e.g.*, some publications assert no effect while others assert an effect) have been flagged as uncertain (supplemental Tables S9 and S10). We observe a good correlation between the A3G and Vif residues binding interface in the model and the residues known to affect binding and/or function. Similarly, we observe that "neutral" residues are scattered on the surface of the proteins and mostly do not localize to the interaction interface.

Third, we assessed whether A3G residues that have been indicated to have structural or functional effect in Vif-mediated A3G degradation localize to the A3G-Vif interface. Comparative structural analyses have determined that Vif-binding A3 domains have negatively charged regions (89) and suggest A3G negatively charged residues at α3, α4, loop 6, loop 8, β3, β4, and β5 are part of the A3G-Vif interface. Notably, negatively charged residues within this region have an average contact frequency of 68% in the integrative modeling ensemble (supplemental Table S9). Structure-guided mutagenesis studies (90) and ubiquitination analyses (91, 92) identified four A3G CTD lysines (297, 301, 303, and 334) that are implicated in Vif-dependent A3G ubiquitination and degradation. In our integrative structure, all of these lysines localize far from the A3G-Vif interface, in the flexible A3G CTD. The large dynamic range of conformations adopted by the A3G CTD may help facilitate ubiquitin transfer.

DISCUSSION

In this work, we present a pipeline to streamline the structure characterization of host–pathogen complexes by using integrative structure modeling based on chemical cross-link data and predicted residue–protein contacts from mutagenesis studies. This pipeline was validated by determining the structure of the A3G-Vif-CRL5-CBFβ complex. By pooling cross-linked peptides from different Vif-containing subcomplexes (Fig. 1), we identified a large number of DSSO-modified peptides, thus improving the coverage of the sequences detected by MS and the number of cross-links for each subunit. We identified 132 high-confidence cross-links, for which we assigned a composite confidence score based on their frequencies and the MS scores. This cross-link dataset, as well as atomic structures of the subunits, and highly reproducible data from mutagenesis studies, enabled us to compute the integrative structure of the entire heptameric A3G-Vif-CRL5-CBFβ complex. We show that including the cross-links composite confidence scores and the mutagenesis data improved the precision of the structural ensembles obtained using integrative structure modeling. In addition, we exemplify how integrative structures can be validated by randomly removing data and resampling to determine if the data types used are consistent with each other and to assess the sources of heterogeneity in the data. This study shows the feasibility of using DSSO-based XL-MS analysis for integrative modeling of structurally heterogeneous host–pathogen protein complexes.

The structure of the full-length A3G-Vif-CRL5-CBFβ complex indicates that Vif predominantly interacts with the A3G NTD. Although the A3G-Vif interaction interface has been characterized in multiple mutagenesis experiments (see supplemental Tables S9 and S10 for extensive review; summarized in supplemental Fig. S13), our modeling combines these biochemical and genetically identified residues with cross-link data to describe the heptameric A3G-Vif-CRL5-CBFβ. The A3G-Vif interface includes regions that have been predicted to be part of the interface by multiple mutagenesis experiments (*i.e.*, A3G residues 126–132, loops α1-β1, β2-α2, and β4-α4; and Vif 40–45 and 22–26) (21, 41, 93), as well as regions that have not previously predicted (*e.g.*, A3G: 30–35, 55–66, 92–101, and 188–195). Importantly, we obtained a model ensemble that satisfies all the input information, including 99% of the input cross-links and the distances implied by the two residue–protein proximity restraints.

Cross-linking data is derived from solution samples that may exhibit structural heterogeneity within the timescale of the experiment. The origin of heterogeneity can be dynamic or static (94–96). Whereas static heterogeneity may arise from different stable conformational states that do not interconvert between each other on the timescale of the experiment, dynamic heterogeneity corresponds to structural fluctuations of the system on the timescale of the experiment. In the case of A3G-Vif-CRL5-CBFβ, it is likely that the cross-linking data reflect the dynamic heterogeneity of the system, given the samples are in solution at 37 °C. While A3G and Cul5 are known to be flexible (78, 80) to allow for the functional ubiquitin transfer (97), our cross-link dataset and integrative model provide additional evidence to support this notion. First, violations of some cross-links by the X-ray structure of Vif-CBC-Cul5$_{NTD}$ and comparative model of Vif-CRL5-CBFβ subcomplexes indicate that the cross-link dataset reflects either a different conformational state or a range of possible structures of the A3G-Vif-CRL5-CBFβ complex. Furthermore, integrative modeling using a rigid representation of Cul5 resulted in violations of a large number of Cul5 intra- and inter-subunit cross-links, despite thorough structural sampling. These violations imply that Cul5 adopts multiple conformations in solution. When a flexible representation of Cul5 was used, the model satisfies 99% of the cross-links. Second, our model indicates that the A3G CTD is not well localized. In integrative modeling, poor localization of a protein or domain implies a lack of information, heterogeneity in the structure, or a combination of both. Modeling using only a fraction of the cross-link dataset did not affect the predicted A3G-Vif

interface or the model precision, indicating that the uncertainty in the output model reflects primarily structural heterogeneity of the sample, not the lack of data. Moreover, the A3G CTD contains seven lysines, five of which reacted with DSSO. In comparison, the A3G NTD contains 13 lysines, ten of which reacted with DSSO. Qualitatively, lysines in both A3G domains are similarly reactive to DSSO and peptides derived from these domains are similarly detected by MS. Consequently, the data indicate differences in localization of the two A3G domains, rather than a difference in their reactivity or detectability in XL-MS experiments. Lastly, we used a Bayesian scoring function (98, 99) to simultaneously model an ensemble of A3G-Vif-CRL5-CBFβ structures and infer additional parameters, such as the uncertainty in the cross-linking data. This Bayesian approach revealed that the computed data uncertainty is consistent with the independently estimated uncertainty of the experimental data (supplemental Table S10), indicating that the ensemble of models appropriately represents the system.

Distinct binding interfaces, corresponding to the different members of the A3 family (*i.e.*, A3G, A3F, or A3H), have been mapped to the N-terminal **α/β** domain of Vif. For example, A3F binds to a site involving Vif residues 11 to 17, 74 to 80, and 171 to 174 (100–102). A3H binding interface involves Vif residues 39, 48, and the 60 to 63 segment (103–105). The normalized contact frequencies indicate that Vif residues involved in binding A3F and A3G have normalized contact frequencies below 60% in the A3G-Vif-CRL5-CBFβ integrative structure. In contrast, residues involved in A3G binding have normalized contact frequencies above 90% (supplemental Table S9). These high normalized contact frequencies are observed for Vif residues used for modeling in the residue–protein binding restraint (40–45) as well as for Vif residues whose binding information was not used for modeling (*i.e.*, Vif residues 22, 26, 30, 53, and 70). These results indicate that our integrative approach identified the binding site on Vif that binds specifically A3G, but not to other A3s.

Reactivation of A3s restriction activity is an attractive HIV-1 treatment strategy. It has been shown that disruption of the A3G-Vif-CRL5-CBFβ complexes as well as inhibition of downstream components, such as the proteasome, can restore antiviral activities and attenuate HIV infectivity (106, 107). Current therapeutic strategies targeting Vif include disrupting the Vif multimerization, A3G-Vif interface, Vif-EloC interface, Vif-CBFβ, and developing A3G up regulators (108, 109). By identifying the larger A3G-Vif interface, we might be able to design a combination strategy that prevents the virus from escaping A3G antiviral activity *via* a single point mutation.

The experimental and computational protocols described here are generally applicable to other difficult to characterize host–pathogen protein complexes. Moreover, this XL-MS and integrative modeling approach can be expanded to include other orthogonal data types, as well as data collected *in vivo*. For example, these XL-MS methods can be applied *in vivo* to infected cells for integrative structure determination of host–pathogen protein assemblies that are difficult to purify, are involved in viral-mediated signaling, localize to special cellular compartments, or are membrane associated. As the field of XL-MS moves toward intracellular applications, the potential for studying host–pathogen complexes, pathways, and networks on an unbiased level becomes more feasible. Complementary to this *in vivo* XL-MS application is the use of *in vivo* genetic interactions obtained using the point mutant epistatic miniarray profile (pE-MAP) platform to derive spatial restraints for integrative modeling (110–112). In such cases, the structure of host–pathogen complexes and how they affect infection can be studied by introducing specific mutations into the pathogenic genome and studying the phenotypic consequences using genetic interaction profiling of relevant host genes. Combining XL-MS and pE-MAP data has proven to improve the precision and accuracy of the models (111).

DATA AVAILABILITY

The mass spectrometry proteomics data are available *via* ProteomeXchange with the dataset identifier PXD025391. Annotated spectra for all interlinked, loop-linked, dead-end, and single peptides can be found on the MS-Viewer application through ProteinProspector (58) (https://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) with the search key 9tjmaqhszr. All the relevant scripts, data, and results are available at GitHub, https://github.com/integrativemodeling/A3G-CRL5-Vif-CBFb. The integrative structures of A3G-Vif-CRL5-Vif-CBFβ have been deposited at PDB-Dev (https://pdb-dev.wwpdb.org/) with the accession code PDBDEV_00000090 and PDBDEV_00000091. The open-source software suite for integrative structure modeling, Integrative Modeling Platform (IMP), is available at https://integrativemodeling.org.

Supplemental data—This article contains supplemental data (5, 7, 17, 20–22, 48, 58, 64, 71, 73–76, 99, 113–177).

Funding and additional information—This work was supported by National Institutes of Health Grants P50AI150476 (to N. J. K, A. S., and J. G.), R01GM083960 (to A. S.), R01GM074830 and R01GM130144 (to L. H.), F32AI127291 (to R.M.K) and Canadian Institutes of Health Research Project Grant PJT-162407 (to L. C.). S. J. K. was supported by KAIST (G04180038), KAIST Advanced Institute for Science-X (N11200065), and National Research Foundation of Korea (2019R1F1A1063373, 2019R1A6A1A10073887, 2020R1A2C1013246, 2020K1A3A7A09080399, and 2020R1A4A3079755). The Krogan Laboratory has received research support from Vir Biotechnology and F. Hoffmann-La Roche. Funding sources had no input or involvement in the conduct of the research, preparation of the manuscript, study design, data collection, data analysis, data interpretation, or in the decision to submit the article for

publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions—I. E., R. M. K., A. S., and N. J. K. conceptualization; R. M. K. and I. E. formal analysis; R. M. K. investigation; I. E., R. M. K., J. V. D., and S. J. K. methodology; L. C., Y. F., H. T., J. G., and L. H. resources; L. H. software; N. J. K., A. S., L. C., J. G., and N. M. C. supervision; R. M. K., I. E., J. V. D., S. J. K., L. C., Y. F., H. T., J. G., L. H., N. J. K., A. S., and N. M. C. writing—original draft.

Conflict of interest—N. J. K. has consulting agreements with the Icahn School of Medicine at Mount Sinai, New York, Maze Therapeutics, and Interline Therapeutics. He is a shareholder in Tenaya Therapeutics, Maze Therapeutics, and Interline Therapeutics.

Abbreviations—The abbreviations used are: A3G, APOBEC3G; CBFβ, core binding factor beta; CTD, C-terminal domain; DSSO, disuccinimidyl sulfoxide; IMP, integrative modeling platform; NTD, N-terminal domain; VCBC, Vif, CBFβ, elongin-B and elongin-C subcomplex; Vif, viral infectivity factor; Vif$_{con}$, consensus Vif; Vif$_{LAI}$, LAI Vif; XL-MS, cross-linking mass spectrometry.

Received April 15, 2021, and in revised from, July 15, 2021 Published, MCPRO Papers in Press, August 11, 2021, https://doi.org/10.1016/j.mcpro.2021.100132

REFERENCES

1. Pushker, R., Mooney, C., Davey, N. E., Jacqué, J.-M., and Shields, D. C. (2013) Marked variability in the extent of protein disorder within and between viral families. *PLoS One* **8**, e60724
2. Xue, B., Williams, R. W., Oldfield, C. J., Goh, G. K.-M., Dunker, A. K., and Uversky, V. N. (2010) Viral disorder or disordered viruses: Do viral proteins possess unique features? *Protein Pept. Lett.* **17**, 932–951
3. Xue, B., Dunker, A. K., and Uversky, V. N. (2012) Orderly order in protein intrinsic disorder distribution: Disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **30**, 137–149
4. Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.-U., Ban, N., Malmström, L., and Aebersold, R. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **337**, 1348–1352
5. Shi, Y., Fernandez-Martinez, J., Tjioe, E., Pellarin, R., Kim, S. J., Williams, R., Schneidman-Duhovny, D., Sali, A., Rout, M. P., and Chait, B. T. (2014) Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell. Proteomics* **13**, 2927–2943
6. Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H. S., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014) Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* **158**, 1123–1135
7. Kim, S. J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B. D., Hogan, J. A., Upla, P., Chemmama, I. E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A. S., *et al*. (2018) Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**, 475–482
8. Gutierrez, C., Chemmama, I. E., Mao, H., Yu, C., Echeverria, I., Block, S. A., Rychnovsky, S. D., Zheng, N., Sali, A., and Huang, L. (2020) Structural dynamics of the human COP9 signalosome revealed by cross-linking mass spectrometry and integrative modeling. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4088–4098
9. Maiolica, A., Cittaro, D., Borsotti, D., Sennels, L., Ciferri, C., Tarricone, C., Musacchio, A., and Rappsilber, J. (2007) Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol. Cell. Proteomics* **6**, 2200–2211
10. O'Reilly, F. J., and Rappsilber, J. (2018) Cross-linking mass spectrometry: Methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* **25**, 1000–1008
11. Rappsilber, J. (2011) The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540
12. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **9**, 1634–1649
13. Iacobucci, C., Piotrowski, C., Aebersold, R., Amaral, B. C., Andrews, P., Bernfur, K., Borchers, C., Brodie, N. I., Bruce, J. E., Cao, Y., and Others. (2019) First community-wide, comparative cross-linking mass spectrometry study. *Anal. Chem.* **91**, 6953–6961
14. Yu, C., and Huang, L. (2018) Cross-linking mass spectrometry: An emerging technology for interactomics and structural biology. *Anal. Chem.* **90**, 144–165
15. Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.-X., Chen, H.-F., Luo, S.-K., Ding, Y.-H., Wang, L.-H., Hao, Z., Xiu, L.-Y., *et al*. (2012) Identification of cross-linked peptides from complex samples. *Nat. Methods* **9**, 904–906
16. Kaake, R. M., Wang, X., Burke, A., Yu, C., Kandur, W., Yang, Y., Novtisky, E. J., Second, T., Duan, J., Kao, A., Guan, S., Vellucci, D., Rychnovsky, S. D., and Huang, L. (2014) A new *in vivo* cross-linking mass spectrometry platform to define protein–protein interactions in living cells. *Mol. Cell. Proteomics* **13**, 3533–3543
17. Kao, A., Chiu, C.-L., Vellucci, D., Yang, Y., Patel, V. R., Guan, S., Randall, A., Baldi, P., Rychnovsky, S. D., and Huang, L. (2011) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell. Proteomics* **10**. M110.002212
18. Sinz, A. (2017) Divide and conquer: Cleavable cross-linkers to study protein conformation and protein–protein interactions. *Anal. Bioanal. Chem.* **409**, 33–44
19. Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the pieces together: Integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244
20. Rout, M. P., and Sali, A. (2019) Principles for integrative structural biology studies. *Cell* **177**, 1384–1403
21. Lavens, D., Peelman, F., Van der Heyden, J., Uyttendaele, I., Catteeuw, D., Verhee, A., Van Schoubroeck, B., Kurth, J., Hallenberger, S., Clayton, R., and Tavernier, J. (2010) Definition of the interacting interfaces of Apobec3G and HIV-1 Vif using MAPPIT mutagenesis analysis. *Nucleic Acids Res.* **38**, 1902–1912
22. Uyttendaele, I., Lavens, D., Catteeuw, D., Lemmens, I., Bovijn, C., Tavernier, J., and Peelman, F. (2012) Random mutagenesis MAPPIT analysis identifies binding sites for Vif and Gag in both cytidine deaminase domains of Apobec3G. *PLoS One* **7**, e44143
23. Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99–103
24. Soros, V. B., Yonemoto, W., and Greene, W. C. (2007) Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H. *PLoS Pathog.* **3**, e15
25. Navarro, F., Bollman, B., Chen, H., König, R., Yu, Q., Chiles, K., and Landau, N. R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333**, 374–386
26. Desimmie, B. A., Delviks-Frankenberrry, K. A., Burdick, R. C., Qi, D., Izumi, T., and Pathak, V. K. (2014) Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J. Mol. Biol.* **426**, 1220–1245
27. Iwatani, Y., Takeuchi, H., Strebel, K., and Levin, J. G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: Correlation with antiviral effect. *J. Virol.* **80**, 5992–6002
28. Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X.-F. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060

29. Harris, R. S., Bishop, K. N., Sheehy, A. M., Craig, H. M., Petersen-Mahrt, S. K., Watt, I. N., Neuberger, M. S., and Malim, M. H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809
30. Jäger, S., Kim, D. Y., Hultquist, J. F., Shindo, K., LaRue, R. S., Kwon, E., Li, M., Anderson, B. D., Yen, L., Stanley, D., Mahon, C., Kane, J., Franks-Skiba, K., Cimermancic, P., Burlingame, A., *et al*. (2011) Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**, 371–375
31. Zhang, W., Du, J., Evans, S. L., Yu, Y., and Yu, X.-F. (2011) T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376–379
32. Chen, G., He, Z., Wang, T., Xu, R., and Yu, X.-F. (2009) A patch of positively charged amino acids surrounding the human immunodeficiency virus type 1 Vif SLVx4Yx9Y motif influences its interaction with APOBEC3G. *J. Virol.* **83**, 8674–8682
33. Shlyakhtenko, L. S., Lushnikov, A. Y., Li, M., Lackey, L., Harris, R. S., and Lyubchenko, Y. L. (2011) Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J. Biol. Chem.* **286**, 3387–3395
34. Morse, M., Huo, R., Feng, Y., Rouzina, I., Chelico, L., and Williams, M. C. (2017) Dimerization regulates both deaminase-dependent and deaminase-independent HIV-1 restriction by APOBEC3G. *Nat. Commun.* **8**, 597
35. Feng, Y., and Chelico, L. (2011) Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. *J. Biol. Chem.* **286**, 11415–11426
36. Stenglein, M. D., Matsuo, H., and Harris, R. S. (2008) Two regions within the amino-terminal half of APOBEC3G cooperate to determine cytoplasmic localization. *J. Virol.* **82**, 9591–9599
37. Smith, H. C., Bennett, R. P., Kizilyer, A., McDougall, W. M., and Prohaska, K. M. (2012) Functions and regulation of the APOBEC family of proteins. *Semin. Cell Dev. Biol.* **23**, 258–268
38. Gooch, B. D., and Cullen, B. R. (2008) Functional domain organization of human APOBEC3G. *Virology* **379**, 118–124
39. Bogerd, H. P., Doehle, B. P., Wiegand, H. L., and Cullen, B. R. (2004) A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3770–3774
40. Huthoff, H., and Malim, M. H. (2007) Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation. *J. Virol.* **81**, 3807–3815
41. Letko, M., Booiman, T., Kootstra, N., Simon, V., and Ooms, M. (2015) Identification of the HIV-1 Vif and human APOBEC3G protein interface. *Cell Rep.* **13**, 1789–1799
42. Schröfelbauer, B., Chen, D., and Landau, N. R. (2004) A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3927–3932
43. Mangeat, B., Turelli, P., Liao, S., and Trono, D. (2004) A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action. *J. Biol. Chem.* **279**, 14481–14483
44. Reingewertz, T. H., Britan-Rosich, E., Rotem-Bamberger, S., Viard, M., Jacobs, A., Miller, A., Lee, J. Y., Hwang, J., Blumenthal, R., Kotler, M., and Friedler, A. (2013) Mapping the Vif-A3G interaction using peptide arrays: A basis for anti-HIV lead peptides. *Bioorg. Med. Chem.* **21**, 3523–3532
45. Haché, G., Shindo, K., Albin, J. S., and Harris, R. S. (2008) Evolution of HIV-1 isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G. *Curr. Biol.* **18**, 819–824
46. Feng, Y., Baig, T. T., Love, R. P., and Chelico, L. (2014) Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front. Microbiol.* **5**, 450
47. Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J. R., McGovern, K. E., Clarke, S. C., Shales, M., Mercenne, G., Pache, L., Li, K., Hernandez, H., Jang, G. M., Roth, S. L., Akiva, E., Marlett, J., *et al*. (2011) Global landscape of HIV-human protein complexes. *Nature* **481**, 365–370
48. Guo, Y., Dong, L., Qiu, X., Wang, Y., Zhang, B., Liu, H., Yu, Y., Zang, Y., Yang, M., and Huang, Z. (2014) Structural basis for hijacking CBF-β and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229–233
49. Jäger, S., Gulbahce, N., Cimermancic, P., Kane, J., He, N., Chou, S., D'Orso, I., Fernandes, J., Jang, G., Frankel, A. D., Alber, T., Zhou, Q., and Krogan, N. J. (2011) Purification and characterization of HIV–human protein complexes. *Methods* **53**, 13–19
50. Kim, D. Y., Kwon, E., Hartley, P. D., Crosby, D. C., Mann, S., Krogan, N. J., and Gross, J. D. (2013) CBFβ stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression. *Mol. Cell* **49**, 632–644
51. Binning, J. M., Smith, A. M., Hultquist, J. F., Craik, C. S., Caretta Cartozo, N., Campbell, M. G., Burton, L., La Greca, F., McGregor, M. J., Ta, H. M., Bartholomeeusen, K., Peterlin, B. M., Krogan, N. J., Sevillano, N., Cheng, Y., *et al*. (2018) Fab-based inhibitors reveal ubiquitin independent functions for HIV Vif neutralization of APOBEC3 restriction factors. *PLoS Pathog.* **14**, e1006830
52. Chelico, L., Pham, P., Calabrese, P., and Goodman, M. F. (2006) APOBEC3G DNA deaminase acts processively 3′→ 5′ on single-stranded DNA. *Nat. Struct. Mol. Biol.* **13**, 392–399
53. Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., *et al*. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920
54. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008) ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536
55. Wang, X., Cimermancic, P., Yu, C., Schweitzer, A., Chopra, N., Engel, J. L., Greenberg, C., Huszagh, A. S., Beck, F., Sakata, E., Yang, Y., Novitsky, E. J., Leitner, A., Nanni, P., Kahraman, A., *et al*. (2017) Molecular details underlying dynamic structures and regulation of the human 26S proteasome. *Mol. Cell. Proteomics* **16**, 840–854
56. Yu, C., Yang, Y., Wang, X., Guan, S., Fang, L., Liu, F., Walters, K. J., Kaiser, P., and Huang. (2016) Characterization of dynamic UbR-proteasome subcomplexes by *in vivo* cross-linking (X) assisted bimolecular tandem affinity purification (XBAP) and label-free quantitation. *Mol. Cell. Proteomics* **15**, 2279–2292
57. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., *et al*. (2019) The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450
58. Baker, P. R., and Chalkley, R. J. (2014) MS-viewer: A web-based spectral viewer for proteomics results. *Mol. Cell. Proteomics* **13**, 1392–1396
59. Ward, A. B., Sali, A., and Wilson, I. A. (2013) Biochemistry. Integrative structural biology. *Science* **339**, 913–915
60. Brünger, A. T., Clore, G. M., Gronenborn, A. M., Saffrich, R., and Nilges, M. (1993) Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation. *Science* **261**, 328–331
61. Hunter, J. D. (2007) Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95
62. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
63. Brademan, D. R., Riley, N. M., Kwiecien, N. W., and Coon, J. J. (2019) Interactive peptide spectral annotator: A versatile web-based tool for proteomic applications. *Mol. Cell. Proteomics* **18**, S193–S201
64. Viswanath, S., Chemmama, I. E., Cimermancic, P., and Sali, A. (2017) Assessing exhaustiveness of stochastic sampling for integrative modeling of macromolecular structures. *Biophys. J.* **113**, 2344–2353
65. Saltzberg, D. J., Viswanath, S., Echeverria, I., Chemmama, I. E., Webb, B., and Sali, A. (2021) Using Integrative Modeling Platform to compute, validate, and archive a model of a protein complex structure. *Protein Sci.* **30**, 250–261
66. Chalkley, R. J., Baker, P. R., Huang, L., Hansen, K. C., Allen, N. P., Rexach, M., and Burlingame, A. L. (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. New developments in protein prospector allow for reliable and comprehensive automatic analysis of large datasets. *Mol. Cell. Proteomics* **4**, 1194–1204
67. Baker, P. R., Trinidad, J. C., and Chalkley, R. J. (2011) Modification site localization scoring integrated into a search engine. *Mol. Cell. Proteomics* **10**. M111.008078
68. Zhang, Y., Wen, Z., Washburn, M. P., and Florens, L. (2010) Refinements to label free proteome quantitation: How to deal with peptides shared by multiple proteins. *Anal. Chem.* **82**, 2272–2281

69. de Vries, S. J., van Dijk, M., and Bonvin, A. M. J. J. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **5**, 883–897
70. Mashiach, E., Schneidman-Duhovny, D., Peri, A., Shavit, Y., Nussinov, R., and Wolfson, H. J. (2010) An integrated suite of fast docking algorithms. *Proteins* **78**, 3197–3204
71. Merkley, E. D., Rysavy, S., Kahraman, A., Hafen, R. P., Daggett, V., and Adkins, J. N. (2014) Distance restraints from crosslinking mass spectrometry: Mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23**, 747–759
72. Webb, B., and Sali, A. (2014) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **47**, 5.6.1–5.6.32
73. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
74. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Rout, M. P., and Sali, A. (2007) Determining the architectures of macromolecular assemblies. *Nature* **450**, 683–694
75. Kouno, T., Luengas, E. M., Shigematsu, M., Shandilya, S. M. D., Zhang, J., Chen, L., Hara, M., Schiffer, C. A., Harris, R. S., and Matsuo, H. (2015) Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. *Nat. Struct. Mol. Biol.* **22**, 485–491
76. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., and Others. (2002) Structure of the Cul1–Rbx1–Skp1–F box Skp2 SCF ubiquitin ligase complex. *Nature* **416**, 703
77. Kwon, Y., Kaake, R. M., Echeverria, I., Suarez, M., Karimian Shamsabadi, M., Stoneham, C., Ramirez, P. W., Kress, J., Singh, R., Sali, A., Krogan, N., Guatelli, J., and Jia, X. (2020) Structural basis of CD4 downregulation by HIV-1 Nef. *Nat. Struct. Mol. Biol.* **27**, 822–828
78. Gorle, S., Pan, Y., Sun, Z., Shlyakhtenko, L. S., Harris, R. S., Lyubchenko, Y. L., and Vuković, L. (2017) Computational model and dynamics of monomeric full-length APOBEC3G. *ACS Cent. Sci.* **3**, 1180–1188
79. [preprint] Pan, Y., Shlyakhtenko, L. S., and Lyubchenko, Y. L. (2019) Insight into dynamics of APOBEC3G protein in complexes with DNA assessed by high speed AFM. *bioRxiv*. https://doi.org/10.1101/581793
80. Liu, J., and Nussinov, R. (2011) Flexible cullins in cullin-RING E3 ligases allosterically regulate ubiquitination. *J. Biol. Chem.* **286**, 40934–40942
81. Onel, M., Sumbul, F., Liu, J., Nussinov, R., and Haliloglu, T. (2017) Cullin neddylation may allosterically tune polyubiquitin chain length and topology. *Biochem. J.* **474**, 781–795
82. Rabut, G., and Peter, M. (2008) Function and regulation of protein neddylation. *EMBO Rep.* **9**, 969–976
83. Calabrese, M. F., Scott, D. C., Duda, D. M., Grace, C. R. R., Kurinov, I., Kriwacki, R. W., and Schulman, B. A. (2011) A RING E3-substrate complex poised for ubiquitin-like protein transfer: Structural insights into cullin-RING ligases. *Nat. Struct. Mol. Biol.* **18**, 947–949
84. Yang, H., Ito, F., Wolfe, A. D., Li, S., Mohammadzadeh, N., Love, R. P., Yan, M., Zirkle, B., Gaba, A., Chelico, L., and Chen, X. S. (2020) Understanding the structural basis of HIV-1 restriction by the full length double-domain APOBEC3G. *Nat. Commun.* **11**, 632
85. Fukuda, H., Li, S., Sardo, L., Smith, J. L., Yamashita, K., Sarca, A. D., Shirakawa, K., Standley, D. M., Takaori-Kondo, A., and Izumi, T. (2019) Structural determinants of the APOBEC3G N-terminal domain for HIV-1 RNA association. *Front. Cell. Infect. Microbiol.* **9**, 129
86. Huthoff, H., Autore, F., Gallois-Montbrun, S., Fraternali, F., and Malim, M. H. (2009) RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS Pathog.* **5**, e1000330
87. Bach, D., Peddi, S., Mangeat, B., Lakkaraju, A., Strub, K., and Trono, D. (2008) Characterization of APOBEC3G binding to 7SL RNA. *Retrovirology* **5**, 54
88. Santa-Marta, M., da Silva, F. A., Fonseca, A. M., and Goncalves, J. (2005) HIV-1 Vif can directly inhibit apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G-mediated cytidine deamination by using a single amino acid interaction and without protein degradation. *J. Biol. Chem.* **280**, 8765–8775
89. Shandilya, S. M. D., Bohn, M.-F., and Schiffer, C. A. (2014) A computational analysis of the structural determinants of APOBEC3's catalytic activity and vulnerability to HIV-1 Vif. *Virology* **471-473**, 105–116
90. Iwatani, Y., Chan, D. S. B., Liu, L., Yoshii, H., Shibata, J., Yamamoto, N., Levin, J. G., Gronenborn, A. M., and Sugiura, W. (2009) HIV-1 Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19539–19544
91. Hüttenhain, R., Xu, J., Burton, L. A., Gordon, D. E., Hultquist, J. F., Johnson, J. R., Satkamp, L., Hiatt, J., Rhee, D. Y., Baek, K., Crosby, D. C., Frankel, A. D., Marson, A., Harper, J. W., Alpi, A. F., *et al*. (2019) ARIH2 is a Vif-dependent regulator of CUL5-mediated APOBEC3G degradation in HIV infection. *Cell Host Microbe* **26**, 86–99.e7
92. Albin, J. S., Anderson, J. S., Johnson, J. R., Harjes, E., Matsuo, H., Krogan, N. J., and Harris, R. S. (2013) Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F. *J. Mol. Biol.* **425**, 1172–1182
93. Letko, M., Silvestri, G., Hahn, B. H., Bibollet-Ruche, F., Gokcumen, O., Simon, V., and Ooms, M. (2013) Vif proteins from diverse primate lentiviral lineages use the same binding site in APOBEC3G. *J. Virol.* **87**, 11861–11871
94. Frauenfelder, H., McMahon, B. H., Austin, R. H., Chu, K., and Groves, J. T. (2001) The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 2370–2374
95. Lu, H. P., Xun, L., and Xie, X. S. (1998) Single-molecule enzymatic dynamics. *Science* **282**, 1877–1882
96. Michalet, X., Weiss, S., and Jäger, M. (2006) Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem. Rev.* **106**, 1785–1813
97. Duda, D. M., Scott, D. C., Calabrese, M. F., Zimmerman, E. S., Zheng, N., and Schulman, B. A. (2011) Structural regulation of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* **21**, 257–264
98. Molnar, K. S., Bonomi, M., Pellarin, R., Clinthorne, G. D., Gonzalez, G., Goldberg, S. D., Goulian, M., Sali, A., and DeGrado, W. F. (2014) Cys-scanning disulfide crosslinking and bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **22**, 1239–1251
99. Rieping, W., Habeck, M., and Nilges, M. (2005) Inferential structure determination. *Science* **309**, 303–306
100. Russell, R. A., and Pathak, V. K. (2007) Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. *J. Virol.* **81**, 8201–8210
101. He, Z., Zhang, W., Chen, G., Xu, R., and Yu, X.-F. (2008) Characterization of conserved motifs in HIV-1 Vif required for APOBEC3G and APOBEC3F interaction. *J. Mol. Biol.* **381**, 1000–1011
102. Dang, Y., Davis, R. W., York, I. A., and Zheng, Y.-H. (2010) Identification of 81LGxGxxIxW89 and 171EDRW174 domains from human immunodeficiency virus type 1 Vif that regulate APOBEC3G and APOBEC3F neutralizing activity. *J. Virol.* **84**, 5741–5750
103. Binka, M., Ooms, M., Steward, M., and Simon, V. (2012) The activity spectrum of Vif from multiple HIV-1 subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. *J. Virol.* **86**, 49–59
104. Ooms, M., Letko, M., Binka, M., and Simon, V. (2013) The resistance of human APOBEC3H to HIV-1 NL4-3 molecular clone is determined by a single amino acid in Vif. *PLoS One* **8**, e57744
105. Refsland, E. W., Hultquist, J. F., Luengas, E. M., Ikeda, T., Shaban, N. M., Law, E. K., Brown, W. L., Reilly, C., Emerman, M., and Harris, R. S. (2014) Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genet.* **10**, e1004761
106. Miyakawa, K., Matsunaga, S., Kanou, K., Matsuzawa, A., Morishita, R., Kudoh, A., Shindo, K., Yokoyama, M., Sato, H., Kimura, H., Tamura, T., Yamamoto, N., Ichijo, H., Takaori-Kondo, A., and Ryo, A. (2015) ASK1 restores the antiviral activity of APOBEC3G by disrupting HIV-1 Vif-mediated counteraction. *Nat. Commun.* **6**, 6945
107. Stanley, D. J., Bartholomeeusen, K., Crosby, D. C., Kim, D. Y., Kwon, E., Yen, L., Cartozo, N. C., Li, M., Jäger, S., Mason-Herr, J., Hayashi, F., Yokoyama, S., Krogan, N. J., Harris, R. S., Peterlin, B. M., *et al*. (2012) Inhibition of a NEDD8 cascade restores restriction of HIV by APOBEC3G. *PLoS Pathog.* **8**, e1003085
108. Bennett, R. P., Salter, J. D., and Smith, H. C. (2018) A new class of antiretroviral enabling innate immunity by protecting APOBEC3 from HIV Vif-dependent degradation. *Trends Mol. Med.* **24**, 507–520
109. Salter, J. D., Morales, G. A., and Smith, H. C. (2014) Structural insights for HIV-1 therapeutic strategies targeting Vif. *Trends Biochem. Sci.* **39**, 373–380
110. Braberg, H., Jin, H., Moehle, E. A., Chan, Y. A., Wang, S., Shales, M., Benschop, J. J., Morris, J. H., Qiu, C., Hu, F., Tang, L. K., Fraser, J. S.,

Holstege, F. C. P., Hieter, P., Guthrie, C., *et al*. (2013) From structure to systems: High-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* **154**, 775–788
111. Braberg, H., Echeverria, I., Bohn, S., Cimermancic, P., Shiver, A., Alexander, R., Xu, J., Shales, M., Dronamraju, R., Jiang, S., Dwivedi, G., Bogdanoff, D., Chaung, K. K., Hüttenhain, R., Wang, S., *et al*. (2020) Genetic interaction mapping informs integrative structure determination of protein complexes. *Science* **370**, aaz4910
112. Gordon, D. E., Watson, A., Roguev, A., Zheng, S., Jang, G. M., Kane, J., Xu, J., Guo, J. Z., Stevenson, E., Swaney, D. L., Franks-Skiba, K., Verschueren, E., Shales, M., Crosby, D. C., Frankel, A. D., *et al*. (2020) A quantitative genetic interaction map of HIV infection. *Mol. Cell* **78**, 197–209.e7
113. Schneidman-Duhovny, D., Pellarin, R., and Sali, A. (2014) Uncertainty in integrative structural modeling. *Curr. Opin. Struct. Biol.* **28**, 96. 10
114. Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012) Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1380–1387
115. Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244
116. Sali, A., Berman, H. M., Schwede, T., Trewhella, J., Kleywegt, G., Burley, S. K., Markley, J., Nakamura, H., Adams, P., Bonvin, A. M. J. J., Chiu, W., Peraro, M. D., Di Maio, F., Ferrin, T. E., Grünewald, K., *et al*. (2015) Outcome of the first wwPDB hybrid/integrative methods task force workshop. *Structure* **23**, 1156–1167
117. Burley, S. K., Kurisu, G., Markley, J. L., Nakamura, H., Velankar, S., Berman, H. M., Sali, A., Schwede, T., and Trewhella, J. (2017) PDB-Dev: a prototype system for depositing integrative/hybrid structural models. *Structure* **25**, 1317–1318
118. Xiao, X., Li, S.-X., Yang, H., and Chen, X. S. (2016) Crystal structures of APOBEC3G N-domain alone and its complex with DNA. *Nat. Commun.* **7**, 12193
119. Webb, B., and Sali, A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **86**, 2.9.1–2.9.37
120. Li, M., Shandilya, S. M. D., Carpenter, M. A., Rathore, A., Brown, W. L., Perkins, A. L., Harki, D. A., Solberg, J., Hook, D. J., Pandey, K. K., Parniak, M. A., Johnson, J. R., Krogan, N. J., Somasundaran, M., Ali, A., Schiffer, C. A., and Harris, R. S. (2012) First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G. *ACS Chem. Biol.* **7**, 506–517
121. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W., and Pavletich, N. P. (2002) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703–709
122. Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B. T., Sali, A., and Rout, M. P. (2007) The molecular architecture of the nuclear pore complex. *Nature* **450**, 695–701
123. Shen, M.-Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524
124. Swendsen, R. H., and Wang, J. S. (1986) Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.* **57**, 2607–2609
125. Vallat, B., Webb, B., Westbrook, J. D., Sali, A., and Berman, H. M. (2018) Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules. *Structure* **26**, 894–904.e2
126. Chodera, J. D. (2016) A simple method for automated equilibration detection in molecular simulations. *J. Chem. Theory Comput.* **12**, 1799–1805
127. McInnes, L., Healy, J., and Astels, S. (2017) hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2**, 205
128. Duda, D. M., Borg, L. A., Scott, D. C., Hunt, H. W., Hammel, M., and Schulman, B. A. (2008) Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134**, 995–1006
129. Sarikas, A., Hartmann, T., and Pan, Z.-Q. (2011) The cullin protein family. *Genome Biol* **12**, 220
130. Mehle, A., Thomas, E. R., Rajendran, K. S., and Gabuzda, D. (2006) A zinc-binding region in Vif binds Cul5 and determines cullin selection. *J. Biol. Chem.* **281**, 17259–17265
131. Chen, K.-M., Harjes, E., Gross, P. J., Fahmy, A., Lu, Y., Shindo, K., Harris, R. S., and Matsuo, H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* **452**, 116–119
132. Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H., and Katahira, M. (2009) Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. *EMBO J.* **28**, 440–451
133. Harjes, E., Gross, P. J., Chen, K.-M., Lu, Y., Shindo, K., Nowarski, R., Gross, J. D., Kotler, M., Harris, R. S., and Matsuo, H. (2009) An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. *J. Mol. Biol.* **389**, 819–832
134. Holden, L. G., Prochnow, C., Chang, Y. P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R. C., Goodman, M. F., and Chen, X. S. (2008) Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121–124
135. Shandilya, S. M. D., Nalam, M. N. L., Nalivaika, E. A., Gross, P. J., Valesano, J. C., Shindo, K., Li, M., Munson, M., Royer, W. E., Harjes, E., Kono, T., Matsuo, H., Harris, R. S., Somasundaran, M., and Schiffer, C. A. (2010) Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. *Structure* **18**, 28–38
136. Lu, X., Zhang, T., Xu, Z., Liu, S., Zhao, B., Lan, W., Wang, C., Ding, J., and Cao, C. (2015) Crystal structure of DNA cytidine deaminase ABOBEC3G catalytic deamination domain suggests a binding mode of full-length enzyme to single-stranded DNA. *J. Biol. Chem.* **290**, 4010–4021
137. Maiti, A., Myint, W., Kanai, T., Delviks-Frankenberry, K., Sierra Rodriguez, C., Pathak, V. K., Schiffer, C. A., and Matsuo, H. (2018) Crystal structure of the catalytic domain of HIV-1 restriction factor APOBEC3G in complex with ssDNA. *Nat. Commun.* **9**, 2460
138. Yan, X., Lan, W., Wang, C., and Cao, C. (2019) Structural investigations on the interactions between cytidine deaminase human APOBEC3G and DNA. *Chem. Asian J.* **14**, 2235–2241
139. Lumpkin, R. J., Baker, R. W., Leschziner, A. E., and Komives, E. A. (2020) Structure and dynamics of the ASB9 CUL-RING E3 ligase. *Nat. Commun.* **11**, 2866
140. Bulliard, Y., Turelli, P., Röhrig, U. F., Zoete, V., Mangeat, B., Michielin, O., and Trono, D. (2009) Functional analysis and structural modeling of human APOBEC3G reveal the role of evolutionarily conserved elements in the inhibition of human immunodeficiency virus type 1 infection and Alu transposition. *J. Virol.* **83**, 12611–12621
141. Rathore, A., Carpenter, M. A., Demir, Ö., Ikeda, T., Li, M., Shaban, N. M., Law, E. K., Anokhin, D., Brown, W. L., Amaro, R. E., and Harris, R. S. (2013) The local dinucleotide preference of APOBEC3G can be altered from 5′-CC to 5′-TC by a single amino acid substitution. *J. Mol. Biol.* **425**, 4442–4454
142. Zhai, C., Ma, L., Zhang, Z., Ding, J., Wang, J., Zhang, Y., Li, X., Guo, F., Yu, L., Zhou, J., and Cen, S. (2017) Identification and characterization of loop7 motif and its role in regulating biological function of human APOBEC3G through molecular modeling and biological assay. *Yao Xue Xue Bao* **7**, 571–582
143. Zhang, W., Du, J., Yu, K., Wang, T., Yong, X., and Yu, X.-F. (2010) Association of potent human antiviral cytidine deaminases with 7SL RNA and viral RNP in HIV-1 virions. *J. Virol.* **84**, 12903–12913
144. Ziegler, S. J., Liu, C., Landau, M., Buzovetsky, O., Desimmie, B. A., Zhao, Q., Sasaki, T., Burdick, R. C., Pathak, V. K., Anderson, K. S., and Xiong, Y. (2018) Insights into DNA substrate selection by APOBEC3G from structural, biochemical, and functional studies. *PLoS One* **13**, e0195048
145. Fang, L., and Landau, N. R. (2007) Analysis of Vif-induced APOBEC3G degradation using an alpha-complementation assay. *Virology* **359**, 162–169
146. Kobayashi, M., Takaori-Kondo, A., Miyauchi, Y., Iwai, K., and Uchiyama, T. (2005) Ubiquitination of APOBEC3G by an HIV-1 Vif-Cullin5-Elongin B-Elongin C complex is essential for Vif function. *J. Biol. Chem.* **280**, 18573–18578
147. Schröfelbauer, B., Senger, T., Manning, G., and Landau, N. R. (2006) Mutational alteration of human immunodeficiency virus type 1 Vif allows for functional interaction with nonhuman primate APOBEC3G. *J. Virol.* **80**, 5984–5991

148. Xu, H., Svarovskaia, E. S., Barr, R., Zhang, Y., Khan, M. A., Strebel, K., and Pathak, V. K. (2004) A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5652–5657
149. Ai, Y., Ma, J., and Wang, X. (2017) Clues for two-step virion infectivity factor regulation by core binding factor beta. *J. Gen. Virol.* **98**, 1113–1121
150. Bergeron, J. R. C., Huthoff, H., Veselkov, D. A., Beavil, R. L., Simpson, P. J., Matthews, S. J., Malim, M. H., and Sanderson, M. R. (2010) The SOCS-box of HIV-1 Vif interacts with ElonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. *PLoS Pathog.* **6**, e1000925
151. Britan-Rosich, E., Nowarski, R., and Kotler, M. (2011) Multifaceted counter-APOBEC3G mechanisms employed by HIV-1 Vif. *J. Mol. Biol.* **410**, 1065–1076
152. Dang, Y., Wang, X., Zhou, T., York, I. A., and Zheng, Y.-H. (2009) Identification of a novel WxSLVK motif in the N terminus of human immunodeficiency virus and simian immunodeficiency virus Vif that is critical for APOBEC3G and APOBEC3F neutralization. *J. Virol.* **83**, 8544–8552
153. Dang, Y., Wang, X., York, I. A., and Zheng, Y.-H. (2010) Identification of a critical T(Q/D/E)x5ADx2(I/L) motif from primate lentivirus Vif proteins that regulate APOBEC3G and APOBEC3F neutralizing activity. *J. Virol.* **84**, 8561–8570
154. DeHart, J. L., Bosque, A., Harris, R. S., and Planelles, V. (2008) Human immunodeficiency virus type 1 Vif induces cell cycle delay via recruitment of the same E3 ubiquitin ligase complex that targets APOBEC3 proteins for degradation. *J. Virol.* **82**, 9265–9272
155. Fujita, M., Sakurai, A., Yoshida, A., Matsumoto, S., Miyaura, M., and Adachi, A. (2002) Subtle mutations in the cysteine region of HIV-1 Vif drastically alter the viral replication phenotype. *Microbes Infect.* **4**, 621–624
156. Guy, B., Geist, M., Dott, K., Spehner, D., Kieny, M. P., and Lecocq, J. P. (1991) A specific inhibitor of cysteine proteases impairs a Vif-dependent modification of human immunodeficiency virus type 1 Env protein. *J. Virol.* **65**, 1325–1331
157. Kataropoulou, A., Bovolenta, C., Belfiore, A., Trabatti, S., Garbelli, A., Porcellini, S., Lupo, R., and Maga, G. (2009) Mutational analysis of the HIV-1 auxiliary protein Vif identifies independent domains important for the physical and functional interaction with HIV-1 reverse transcriptase. *Nucleic Acids Res.* **37**, 3660–3669
158. Ma, X. Y., Sova, P., Chao, W., and Volsky, D. J. (1994) Cysteine residues in the Vif protein of human immunodeficiency virus type 1 are essential for viral infectivity. *J. Virol.* **68**, 1714–1720
159. Matsui, Y., Shindo, K., Nagata, K., Yoshinaga, N., Shirakawa, K., Kobayashi, M., and Takaori-Kondo, A. (2016) Core binding factor β protects HIV, Type 1 accessory protein viral infectivity factor from MDM2-mediated degradation. *J. Biol. Chem.* **291**, 24892–24899
160. Mehle, A., Wilson, H., Zhang, C., Brazier, A. J., McPike, M., Pery, E., and Gabuzda, D. (2007) Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *J. Virol.* **81**, 13235–13241
161. Miyagi, E., Kao, S., Yedavalli, V., and Strebel, K. (2014) CBFβ enhances de novo protein biosynthesis of its binding partners HIV-1 Vif and RUNX1 and potentiates the Vif-induced degradation of APOBEC3G. *J. Virol.* **88**, 4839–4852
162. Mulder, L. C. F., Ooms, M., Majdak, S., Smedresman, J., Linscheid, C., Harari, A., Kunz, A., and Simon, V. (2010) Moderate influence of human APOBEC3F on HIV-1 replication in primary lymphocytes. *J. Virol.* **84**, 9613–9617
163. Nakashima, M., Ode, H., Kawamura, T., Kitamura, S., Naganawa, Y., Awazu, H., Tsuzuki, S., Matsuoka, K., Nemoto, M., Hachiya, A., Sugiura, W., Yokomaku, Y., Watanabe, N., and Iwatani, Y. (2015) Structural insights into HIV-1 Vif-APOBEC3F interaction. *J. Virol.* **90**, 1034–1047
164. Simon, V., Zennou, V., Murray, D., Huang, Y., Ho, D. D., and Bieniasz, P. D. (2005) Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog* **1**, e6
165. Stanley, B. J., Ehrlich, E. S., Short, L., Yu, Y., Xiao, Z., Yu, X.-F., and Xiong, Y. (2008) Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *J. Virol.* **82**, 8656–8663
166. Tian, C., Yu, X., Zhang, W., Wang, T., Xu, R., and Yu, X.-F. (2006) Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F. *J. Virol.* **80**, 3112–3115
167. Walker, R. C., JrKhan, M. A., Kao, S., Goila-Gaur, R., Miyagi, E., and Strebel, K. (2010) Identification of dominant negative human immunodeficiency virus type 1 Vif mutants that interfere with the functional inactivation of APOBEC3G by virus-encoded Vif. *J. Virol.* **84**, 5201–5211
168. Wang, H., Liu, B., Liu, X., Li, Z., Yu, X.-F., and Zhang, W. (2014) Identification of HIV-1 Vif regions required for CBF-β interaction and APOBEC3 suppression. *PLoS One* **9**, e95738
169. Wichroski, M. J., Ichiyama, K., and Rana, T. M. (2005) Analysis of HIV-1 viral infectivity factor-mediated proteasome-dependent depletion of APOBEC3G: correlating function and subcellular localization. *J. Biol. Chem.* **280**, 8387–8396
170. Xiao, Z., Xiong, Y., Zhang, W., Tan, L., Ehrlich, E., Guo, D., and Yu, X.-F. (2007) Characterization of a novel Cullin5 binding domain in HIV-1 Vif. *J. Mol. Biol.* **373**, 541–550
171. Yamashita, T., Kamada, K., Hatcho, K., Adachi, A., and Nomaguchi, M. (2008) Identification of amino acid residues in HIV-1 Vif critical for binding and exclusion of APOBEC3G/F. *Microbes Infect.* **10**, 1142–1149
172. Yamashita, T., Nomaguchi, M., Miyake, A., Uchiyama, T., and Adachi, A. (2010) Status of APOBEC3G/F in cells and progeny virions modulated by Vif determines HIV-1 infectivity. *Microbes Infect.* **12**, 166–171
173. Yang, X., Goncalves, J., and Gabuzda, D. (1996) Phosphorylation of Vif and its role in HIV-1 replication. *J. Biol. Chem.* **271**, 10121–10129
174. Yang, X., and Gabuzda, D. (1998) Mitogen-activated protein kinase phosphorylates and regulates the HIV-1 Vif protein. *J. Biol. Chem.* **273**, 29879–29887
175. Yu, Y., Xiao, Z., Ehrlich, E. S., Yu, X., and Yu, X.-F. (2004) Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev.* **18**, 2867–2872
176. Zhang, H., Pomerantz, R. J., Dornadula, G., and Sun, Y. (2000) Human immunodeficiency virus type 1 Vif protein is an integral component of an mRNP complex of viral RNA and could be involved in the viral RNA folding and packaging process. *J. Virol.* **74**, 8252–8261
177. Zhang, W., Chen, G., Niewiadomska, A. M., Xu, R., and Yu, X.-F. (2008) Distinct determinants in HIV-1 Vif and human APOBEC3 proteins are required for the suppression of diverse host anti-viral proteins. *PLoS One* **3**, e3963