Elucidation of New Binding Interactions with the Human Tsg101 Protein Using Modified HIV-1 Gag-p6 Derived Peptide Ligands

Sung-Eun Kim,[†] Fa Liu,[†] Young Jun Im,[‡] Andrew G. Stephen,[§] Matthew J. Fivash,[∥] Abdul A. Waheed,[⊥] Eric O. Freed,[⊥] Robert J. Fisher,[§] James H. Hurley,[‡] and Terrence R. Burke*,[†]

[†]Chemical Biology Laboratory, Molecular, Discovery Program, CCR, NCI-Frederick, Frederick, Maryland 21702, United States
[‡]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney and [§]Protein Chemistry Laboratory, Advanced Technology Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702, United States
[∥]Data Management Systems, Inc., NCI-Frederick, Frederick, Maryland 21702, United States
[⊥]HIV Drug Resistance Program, CCR, NCI-Frederick, Frederick, Maryland 21702, United States

S *Supporting Information*

ABSTRACT: Targeting protein–protein interactions is gaining greater recognition as an attractive approach to therapeutic development. An example of this may be found with the human cellular protein encoded by the tumor susceptibility gene 101 (Tsg101), where interaction with the p6 C-terminal domain of the nascent viral Gag protein is required for HIV-1 particle budding and release. This association of Gag with Tsg101 is highly dependent on a "Pro-Thr-Ala-Pro" ("PTAP") peptide sequence within the p6 protein. Although p6-derived peptides offer potential starting points for developing Tsg101-binding inhibitors, the affinities of canonical peptides are outside the useful range (K_d values greater than 50 μM). Reported herein are crystal structures of Tsg101 in complex with two structurally modified PTAP-derived peptides. These data define new regions of ligand interaction not previously identified with canonical peptide sequences. This information could be highly useful in the design of Tsg101-binding antagonists.
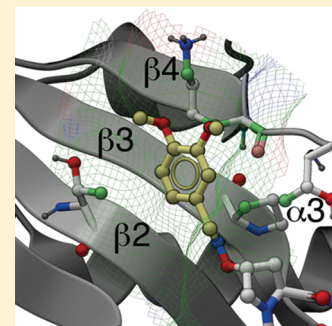
KEYWORDS: Protein–protein interactions, Tsg101, X-ray crystal structure, peptide analogues

H istorically, large components of therapeutic development have been directed at cell membrane receptors or at enzymes, in the latter case where advantage can be taken of well-defined substrate binding clefts or catalytic mechanisms.[1] Recently, targeting protein–protein interactions (PPIs) has gained greater recognition.[2] Because PPIs typically involve contacts over extended and relatively flat surfaces, it may seem counterintuitive that small molecules or peptides could function as effective binding antagonists. However, significant components of total binding PPI energy are often derived from a limited number of protein residues localized in well-defined "hot spot" regions.[3] This allows the possibility of effective overall PPI inhibition by relatively small agents, which can take advantage of high affinity interactions within these hot spot regions.[4] For this reason, development of PPI inhibitors has become an extremely active area of research that is beginning to yield clinically relevant agents.[5,6]

Acquired immunodeficiency syndrome (AIDS) is a devastating disease caused by the human immunodeficiency virus (HIV-1), which after nearly 30 years of research has a total of 25 FDA-approved drugs marketed for its treatment.[7] These agents are either directed against key enzymes in the viral life cycle, such as protease, reverse transcriptase, and integrase, or inhibit viral entry. As typified recently by integrase, the development of drug-resistant mutant enzymes provides a sense of urgency to the continued search for new modes of inhibiting HIV-1 replication.[8]

Because the efficient production of a new virus requires specific interactions between viral and human proteins, pharmacological inhibition of appropriate PPIs represents potentially attractive new opportunities for anti-HIV-1 therapeutic development.[9] One recognized target of this sort is the human cellular protein encoded by the tumor susceptibility gene 101 (Tsg101), which is a component of the host endosomal sorting pathway, whose interaction with the p6 C-terminal domain of the nascent viral Gag protein is required for viral assembly and budding.[10−13] Binding of Gag to the ubiquitin E2 variant (UEV) domain of Tsg101 is dependent on a Pro-Thr-Ala-Pro ("P-T-A-P") motif with the p6 region, and blocking this interaction results in antiviral effects.[14−16] The importance of the relatively confined P-T-A-P sequence to the overall Tsg101–Gag interaction indicates that this might be "hot spot" in nature and amenable to inhibitor development. Consistent with this notion, a previous NMR solution structure of the p6-derived peptide "P^1-E^2-P^3-T^4-A^5-P^6-P^7-E^8-E^9" bound to Tsg101 shows that the peptide binds in a groove, with the P^3-T^4-A^5-P^6 residues making the greatest contact.[17] Although this sequence could serve as an initial starting point for inhibitor development, its affinity is outside the useful range (K_d > 50 μM), and its binding interactions as
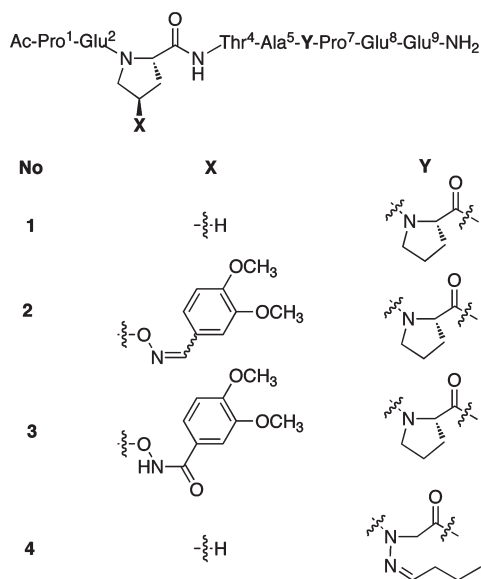
Figure 1. Structures of peptides discussed in the text.

indicated by the NMR solution structure are not sufficiently well-defined to allow true "structure-based design". Therefore, we undertook an empirical approach to ligand development by making structural alterations at each residue of the peptide. This was done in an unbiased fashion using noncoded amino acid derivatives that were synthesized by oxime-based postsolid phase diversification. The protocol involved the insertion of aminooxy groups at each residue of the parent peptide 1 (Figure 1). These then served as "handles" for functional group elaboration by oxime ligation using libraries of aldehydes.[18−20] It was found that while modifications within the region "T[4]-A[5]-P[6]" adversely affected binding, up to 20-fold enhancement of affinity could be incurred by introduction of aromatic functionality at the P[3] position.[20] This was unexpected, since based on the previously reported NMR solution structure, the P[3] residue serves a more minor function by interacting with a shallow pocket on the protein surface.[17]

As reported herein, to understand the structural basis for the observed affinity enhancement, we solved the Tsg101 cocrystal structure of the two most potent analogues (2 and 3, Figure 1), which compliments our recent report of the crystal structure of 1 bound to Tsg101.[21] Prior to undertaking this work, the only structural data related to Tsg101−peptide interactions were derived from NMR solution studies.[16,17] In our current study, we find that the peptide backbones of 2 and 3 are superimposable on that of the parent 1.[20] Peptides 2 and 3 contain 3,4-dimethoxybenzyl substituents on the (4R)-aminooxy handles of their P[3] pyrrolidine rings (linkage via oxime and amide bonds, respectively). The cocrystal structures clarify the basis for the affinity enhancement incurred by functionalization at the P[3] position. The new binding interactions are clearly defined and extend significantly beyond what would be possible from the unmodified parent P[3] residue (Figure 2). The dimethoxybenzyl group of 2 makes van der Waals contacts with nonpolar portions of the side chains of T-56, P-71, and K-90, while the oxime linkage contacts the side chain of T-58. The 3-methoxy group accepts a hydrogen bond from a water molecule that is, in turn, bound to the carbonyl oxygen of P-91. However, no direct hydrogen bonds are formed with the Tsg101 protein. The

aminooxy linkage in 3 positions its dimethoxybenzyl group differently than in 2, such that the benzyl rings are rotated approximately 75° with respect to one another, and the corresponding atomic positions are separated by from 2 to 10 Å. The T-56 residue is the only common point of contact shared with 2. Unique contacts are made with the aliphatic portion of K-36 and the Cα of G-57. The carbonyl oxygen of the aminooxy linkage makes a hydrogen bond with a water molecule, but no direct hydrogen bonds are made with the protein. A common theme in these two structures is that the dimethoxybenzyl groups interact face-on with a broad and relatively flat and nonpolar surface formed by the exposed face of the Tsg101 system of antiparallel β-sheets (Figure 2).

The data above suggest that the newly identified binding surface could accommodate additional Tsg101 binding interactions through the use of more extended P[3] functionalities. To explore this possibility, we prepared several polycyclic aryl-containing oximes (5−7, Table 1) that were intended to potentially offer greater interactions with the binding surface (Figure 3).[22−24] Examination of IC$_{50}$ values determined using a fluorescence anisotropy (FA) assay that measured the ability of peptides to compete with an FITC-labeled variant of peptide 2 (see the Supporting Information) showed that the binding affinities of peptides 5−7 were up to 5-fold higher than that of the parent 2. These results indicate that significant latitude exists in the functionality that is compatible with binding in this region.

A more relevant indication of affinity was obtained by measuring the ability of peptides to inhibit the binding of Tsg101 to p6 protein. For this purpose, surface plasmon resonance (SPR) experiments were performed using GST glutathione S-transferase (GST) or GST-p6 captured on anti-GST antibody chips, with Tsg101 protein in solution either alone or in the presence of increasing concentrations of peptides 2 or 6 (see the Supporting Information). From the amount of Tsg101 protein bound to the p6 protein, IC$_{50}$ values were calculated to be 72 and 12 μM for peptides 2 and 6, respectively. Consistent with the FA competition assays, peptide 6 exhibits an approximate 5−6-fold higher Tsg101-binding affinity than parent peptide 2.

While the interactions of P[3] described above are limited primarily to the protein surface, the previous NMR solution structure of the Tsg101-bound P-E-P-T-A-P-P-E-E nonapeptide indicated that the P[6] residue binds in a pocket wedged between the aromatic rings of Y-63 and Y-68. This interaction is reminiscent of the recognition of polyproline sequences by WW and SH3 domains.[17] Our current cocrystals add significant new understanding regarding this pocket and its potential utility for exploiting more extensive binding interactions. What emerges from our new data is that binding of the P[6] pyrrolidine ring between the Y-68 and the F-142 residues shown by the NMR structure merely defines the upper regions of a much deeper pocket. The lower walls of this pocket form a cylindrical cavity bounded by Y-63 and P-139. The depth of the pocket is determined by the V-61 residue, whose side chain isopropyl group defines its lower boundary (Figure 4). The geometry of this pocket is consistent with structure−activity relationship (SAR) data that we had previously observed. For example, introduction of a variety of oxime derivatives from the 4-position of the P[6] pyrrolidine ring was found to uniformly abrogate binding affinity.[20] It is now evident that in each case the added functionality was either too wide or too long to be accommodated within the pocket. In another study, we examined whether
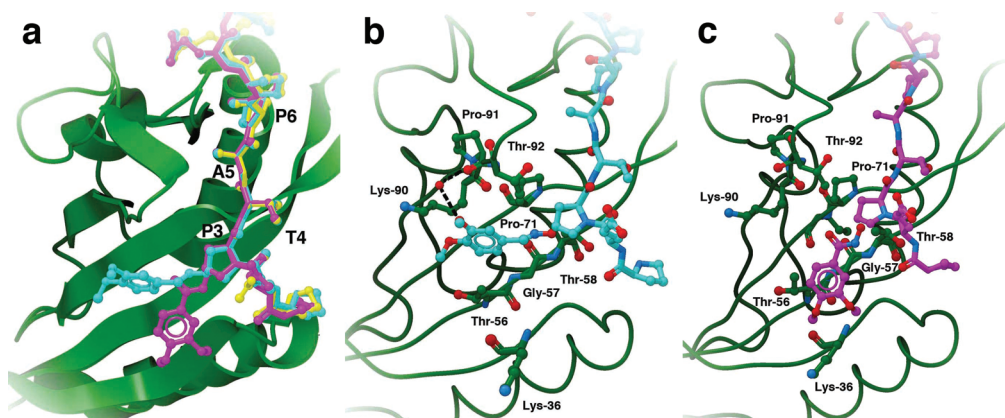
Figure 2. Crystal structures of peptides bound to Tsg101. (a) Superposition of wild-type peptide **1** (yellow; PDB no. 3OBU), peptide **2** (cyan), and peptide **3** (magenta). (b) Binding of peptide **2** showing key residues interacting with the 3,4-dimethoxybenzyl oxime group of Pro-3. (c) Binding of peptide **3** showing key residues interacting with the 3,4-dimethoxybenzyoylcarboxamide group of Pro-3.
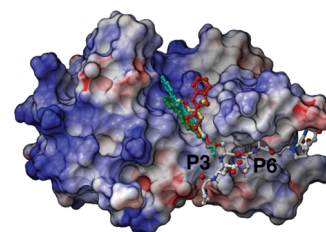
**Table 1. Tsg101 Binding Affinities of P³ Polycyclic Oxime-Containing Peptides**

Ac-Pro¹-Glu²-Thr⁴-Ala⁵-Pro⁶-Pro⁷-Glu⁸-Glu⁹-NH₂
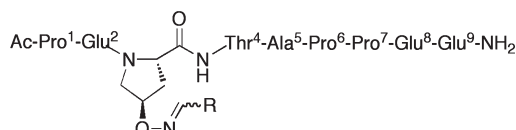
No	R	K_i (µM) ± S.E.[a]
2	–OCH₃ / OCH₃	8.7 ± 3.5
5		5.4 ± 2.2
6		2.7 ± 1.1
7		2.6 ± 1.0

[a] Results of FA competition assays performed as described in the Supporting Information. [b] Structure as indicated in Figure 1.

Figure 3. Examples of potential oxime binding interactions in the newly identified P³ binding region of peptides **5** (cyan), **6** (red), and **7** (green) as compared with parent **2** (yellow). Refer to refs 22 and 23.
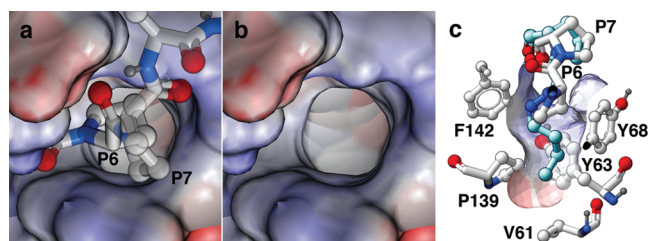
Figure 4. Pro-6 binding pocket from the Tsg101−**2** cocrystal structure rendered as an electrostatic surface potential. (a) Looking down into the pocket showing the positioning of the Pro-6 and Pro-7 residues. (b) The same view as in a following deletion of peptide **2** highlighting the cylindrical nature of the pocket. (c) Cutaway side view showing the electrostatic surface of the pocket with contact residues indicated. For clarity, only the Pro-6 and Pro-7 residues of the peptide are depicted (white). Shown in cyan is the calculated interaction of the butylhydrazone side chain of peptide **4**.

N-substituted glycine (NSG) type structures could serve as effective replacements of the P⁶ residue.[25] This work was based on the previous observation that NSGs can function as high affinity Pro surrogates by more completely filling binding pockets than is possible by the pyrrolidine ring of the parent Pro residue.[26,27] The highest affinity achieved in our earlier study (approximately 5-fold enhancement relative to the reference P⁶-containing parent)[26] was with the butylhydrazone analogue as depicted in peptide **4** (Figure 1).[28] By overlaying the "CH₃−CH₂−CH₂−C═N−" side chain of this NSG mimetic onto the P⁶ pyrrolidine ring of parent peptide **2**, it is now evident that the n-butyl side chain extends nicely into the newly defined P⁶

binding pocket (Figure 4). Therefore, the new crystal data of our current study both clarify previous empirically-derived binding data and provide guidance for the design of new analogues.

In conclusion, we report herein crystal data of ligands complexed to the UEV domain of the human Tsg101 protein that identify new binding interactions originating from the critical P³ and P⁶ residues of the canonical PTAP recognition motif. The P³ proximal region is comprised of an extensive surface topology. As exemplified by several analogues prepared as part of the current study, binding in this region can be exploited by the use of aromatic or polycyclic functionalities. In contrast, the P⁶ binding

pocket is narrow, deep, and well-defined and may provide opportunities for enhancement of both binding affinity and selectivity. This information could be useful in the design of higher affinity Tsg101 binding antagonists.

■ ASSOCIATED CONTENT

ⓢ Supporting Information. Statistics of data collection and crystallographic refinement and 1.8 Å 2Fo-Fc electron density maps for peptides **2** and **3** complexed to Tsg101 protein; procedures for determination of Tsg101-binding affinities of peptides **2** and **5**−**7** using FA competition assays and for peptides **2** and **6** using SPR assays and synthetic protocols for peptides **5**−**7**. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The coordinates of the Tsg101 UEV in complex with peptides **2** and **3** have been deposited with the RCSB Protein Data Bank under the accessions codes of 3P9H (peptide **2**) and 3P9G (peptide **3**).

■ AUTHOR INFORMATION

Corresponding Author
*Tel: 301-846-5906. Fax: 301-846-6033. E-mail: tburke@helix.nih.gov.

■ DISCLOSURE

■ REFERENCES

(1) Drews, J. Drug discovery: A historical perspective. *Science* **2000**, *287*, 1960–1964.

(2) Betzi, S.; Guerlesquin, F.; Morelli, X. Protein-protein interaction inhibition (2P2I): Fewer and fewer undruggable targets. *Comb. Chem. High Throughput Screening* **2009**, *12*, 968–983.

(3) Clackson, T.; Wells, J. A hot spot of binding energy in a hormone-receptor interface. *Science* **1995**, *267*, 383–386.

(4) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001–1009.

(5) Gerrard, J. A.; Hutton, C. A.; Perugini, M. A. Inhibiting protein-protein interactions as an emerging paradigm for drug discovery. *Mini-Rev. Med. Chem.* **2007**, *7*, 151–157.

(6) Zinzalla, G.; Thurston, D. E. Targeting protein-protein interactions for therapeutic intervention: A challenge for the future. *Future Med. Chem.* **2009**, *1*, 65–93.

(7) De Clercq, E. Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *Int. J. Antimicrob. Agents* **2009**, *33*, 307–320.

(8) Hatano, H.; Lampiris, H.; Fransen, S.; Gupta, S.; Huang, W.; Hoh, R.; Martin, J. N.; Lalezari, J.; Bangsberg, D.; Petropoulos, C.; Deeks, S. G. Evolution of integrase resistance during failure of integrase inhibitor-based antiretroviral therapy. *J. Acquired Immune Defic. Syndr.* **2010**, *54*, 389–393.

(9) Rice, A. P.; Sutton, R. E. Targeting protein-protein interactions for HIV therapeutics. *Future HIV Ther.* **2007**, *1*, 369–385.

(10) Huang, M.; Orenstein, J.; Martin, M.; Freed, E. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J. Virol.* **1995**, *69*, 6810–6818.

(11) VerPlank, L.; Bouamr, F.; LaGrassa, T. J.; Agresta, B.; Kikonyogo, A.; Leis, J.; Carter, C. A. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55-(Gag). *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7724–7729.

(12) Garrus, J. E.; von Schwedler, U. K.; Pornillos, O. W.; Morham, S. G.; Zavitz, K. H.; Wang, H. E.; Wettstein, D. A.; Stray, K. M.; Cote, M.; Rich, R. L.; Myszka, D. G.; Sundquist, W. I. Tsg101 and the Vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **2001**, *107*, 55–65.

(13) Martin-Serrano, J.; Zang, T.; Bieniasz, P. D. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat. Med.* **2001**, *7*, 1313–1319.

(14) Freed, E. O. The HIV-TSG101 interface: Recent advances in a budding field. *Trends Microbiol.* **2003**, *11*, 56–59.

(15) Chen, H.; Liu, X.; Li, Z.; Zhan, P.; De Clercq, E. TSG101: A novel anti-HIV-1 drug target. *Curr. Med. Chem.* **2010**, *17*, 750–758.

(16) Demirov, D. G.; Ono, A.; Orenstein, J. M.; Freed, E. O. Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 955–960.

(17) Pornillos, O. W.; Alam, S. L.; Davis, D. R.; Sundquist, W. I. Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nat. Struct. Biol.* **2002**, *9*, 812–817.

(18) Liu, F.; Thomas, J.; Burke, T. R., Jr. Synthesis of a homologous series of side-chain-extended orthogonally protected aminooxy-containing amino acids. *Synthesis* **2008**, 2432–2438.

(19) Liu, F.; Stephen, A. G.; Fisher, R. J.; Burke, T. R., Jr. Protected aminooxyprolines for expedited library synthesis: Application to Tsg101-directed proline-oxime containing peptides. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1096–1101.

(20) Liu, F.; Stephen, A. G.; Waheed, A. A.; Aman, M. J.; Freed, E. O.; Fisher, R. J.; Burke, T. R., Jr. SAR by oxime-containing peptide libraries: Application to Tsg101 ligand optimization. *Chembiochem* **2008**, *9*, 2000–2004.

(21) Im, Y. J.; Kuo, L.; Ren, X.; Burgos, P. V.; Zhao, X.-Z.; Liu, F.; Burke, T. R., Jr.; Bonifacino, J. S.; Freed, E. O.; Hurley, J. H. Crystallographic and functional analysis of the ESCRT-I /HIV-1 Gag PTAP interaction. *Structure (Cambridge, MA)* **2010**, *18*, 1536–1547.

(22) Rigid receptor docking was performed according to standard protocols using ICM Chemist Pro Software v 3.7-1f/MacOSX; MolSoft LLC: La Jolla, CA; https://www.molsoft.com/icm-chemist-pro.html.

(23) Totrov, M.; Abagyan, R. Protein-ligand docking as an energy optimization problem. In *Drug-Receptor Thermodynamics: Introduction and Applications*; Raffa, R. B., Ed.; John Wiley & Sons, Inc.: New York, 2001; pp 603−624.

(24) New peptide oximes were prepared by procedures previously reported in ref 15. Experimental specifics are included in the Supporting Information.

(25) Liu, F.; Stephen, A. G.; Adamson, C. S.; Gousset, K.; Aman, M. J.; Freed, E. O.; Fisher, R. J.; Burke, T. R., Jr. Hydrazone- and hydrazide-containing N-substituted glycines as peptoid surrogates for expedited library synthesis: Application to the preparation of Tsg101-directed HIV-1 budding antagonists. *Org. Lett.* **2006**, *8*, 5165–5168.

(26) Nguyen, J. T.; Turck, C. W.; Cohen, F. E.; Zuckermann, R. N.; Lim, W. A. Exploiting the basis of proline recoginition by SH3 and WW domains: Design of N-substituted inhibitors. *Science* **1998**, *282*, 2088–2091.

(27) Nguyen, J. T.; Porter, M.; Amoui, M.; Miller, W. T.; Zuckermann, R. N.; Lim, W. A. Improving SH3 domain ligand selectivity using a non-natural scaffold. *Chem. Biol.* **2000**, 7, 463–473.

(28) Direct fluoresence anisotropy binding assays reported in refs 20 and 25 were conducted with peptides having FITC groups appended via N-terminal 6-aminohexanoyl chains.