STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 75 (2019)

Supporting information for article:

Structural analysis of free and liganded forms of the Fab fragment of a high-affinity anti-cocaine antibody, h2E2

Kemin Tan, Min Zhou, Angela J. Ahrendt, Norma E. C. Duke, Nassif Tabaja, William J. Ball, Terence L. Kirley, Andrew B. Norman, Andrzej Joachimiak, Marianne Schiffer, Rosemarie Wilton and P. Raj Pokkuluri

S1. Materials and Methods

Primers used for amplification of the pASK88 expression vector and h2E2 Fab domain sequences for seamless cloning are listed in Table S1. Primers used for sequence analysis of the Fab region of the expression vector are given in Table S2. Table S3 contains sequences of the synthetic DNA used to construct the h2E2 Fab. The full sequence of the pASK88_h2E2_Fab expression vector is provided in Table S4.

Table S1 Primers used for vector construction

Primer	Sequence (5'-3')	
pASK88_FW	TAATAAGCTTGACCTGTGAAGTG	For amplification of vector for insertion of Lambda2C gBlock
pASK88_RV	TTTGATCTCGAGCTTGGTCCC	Rev amplification of vector for insertion of Lambda2C gBlock
88CenterF	CCTCCACCAAGGGCCCATCGG	For amplification of vector for insertion of VH domain
88Omp-R	GGCCTGCGCTACGGTAGCG	Rev amplification of vector for insertion of VH domain
882E2VHF	TACCGTAGCGCAGGCCGAAGTGCAGTTGGTTGAATCAGGTGG	For amplification of VH domain from synthetic DNA
882E2VHR	GGCCCTTGGTGGAGGCGCTACTCACCGTCACTAATGTCCC	Rev amplification of VH domain from synthetic DNA
88CenterR	GGCTTTTGTCACAGGGGTAAACAGTAACGG	For amplification of vector for insertion of VL domain
88LamF	GGCCAGCCGAAAGCCGCC	Rev amplification of vector for insertion of VL domain
882E2VLF	CCCTGTGACAAAAGCCCAGGCAGTAGTTATTCAGGAAAGCGC	For amplification of VL domain from synthetic DNA
88Lam2E2VLR	CGGCTTTCGGCTGGCCCAATACTGTAACCTTCGTGCCACCTCC	Rev amplification of VL domain from synthetic DNA

Table S2 Primers used for sequencing

Primer	Sequence (5'-3')	
pASK88P1F	CGAATGGCCAGATGATTAATTCC	Hybridizes in promoter region to read through heavy chain (For)
pASK88P2R	ACAGTAACGGTAAGAGTGCCA	Hybridizes in PhoA leader peptide to read through heavy chain (Rev)
pASK88P3F	GGCACCCAGACCTACATCT	Hybridizes in heavy chain C1 to read through light chain (For)
pASK88P4R	CTTAATGCGCCGCTACAGG	Hybridizes in f1 ori to read through light chain (Rev)

Table S3 Sequences of synthesized DNA fragments used in this study. The lambda2C gBlock has overlaps for seamless cloning (lower case). The h2E2 VH and VL domains (upper case) were synthesized as an scFv with NcoI and NotI sites (underlined) for cloning into pET22B (R. Wilton, unpublished work). Synthetic DNA sequences were expression optimized for *E. coli.*

Fragment	Sequence (5'-3')
huLambd2C gBlock	caagctcgagatcaaaGGCCAGCCGAAAGCCGCCCCCTCCGTCACCCTGTTCCCTCCCTCCTCTGAAGAACTGCAGGCTAACAAAGCGACCCTGGTTTG TTTAATTTCTGACTTCTACCCTGGTGCCGTCACCGTAGCATGGAAAGCCGATTCATCCCCCGTAAAAGCCGGAGTGGAGACCACGACCCCCTCAAAACA GTCCAATAACAAATATGCCGCCTCTTCCTATCTCAGCCTTACCCCCGAACAATGGAAATCACACCGTTCTTATTCCTGCCAAGTCACCCACGAAGGTTC CACCGTTGAAAAAACCGTAGCCCCAACCGAATGCTCCtaataagcttgacctg
h2E2 V_H and V_L domains (formatted as an scFv)	ccatggctGAAGTGCAGTTGGTTGAATCAGGTGGGGGTTTAGTACAGCCGGGTGGTAGTTTACGTTTGTCATGTGCGGCATCAGGTTTTATTTTTAGTA GTGATTGGATGAATTGGGTACGTCAAGCACCGGGAAAAGGATTAGAATGGGTGGCGAATATTAATCAAGATGGTTCAGAAAAATATTATGTGGATTCAG TTAAAGGTCGTTTTACAATCAGCCGTGACAACGCACAAAATAGCTTATACTTACAAATGAACAGTTTACGGGCAGAAGACACAGCAGTATATTATTGTG CAAAGGAATTAGGGCCGTGGGGGCAAGGGACATTAGTGACGGTGAGTAGCgggggagggggcagcggcggtggtggttcgggagggggaggttcgacaC AGGCAGTAGTTATTCAGGAAAGCGCACTCACGACATCTCCGGGGGGGACGGTTATTCTCACTTGCCGCAGCAGTACAGGAACGATTACGACTTCTAACT ATGCAAATTGGGTCCAGAAAAAACCGAATCATGTGTTTACGGGTTTAATTGGGGCAACGAGCATTCGCGCGCCGGGAGTGCCGGTACGTTTTAGCGGGT TTCTTATTGGTGGAAAGGCAGCATTAACTATTACAGGAGCGCAAACCGAAGATGATGCTATGTATTTTTGCGCGTTATGGTATAACACACACTATGTTT TTGGAGGTGGCACGAAGGTTACAGTATTGGGGCAAgcggccgc

Table S4 Full sequence of the h2E2 rFab expression vector

Plasmid	Sequence (5'-3')
pASK88_h2E2_Fab	ACCCGACACCATCGAATGGCCAGATGATTAATTCCTAATTTTTGTTGACACTCTATCATTGATAGAGTTATTTTACCACT CCCTATCAGTGATAGAGAAAAGTGAAATGAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGAAAAAGACAGC TATCGCGATTGCAGTGGCACTGGCTGGTTTcgcTACCGTAGCGCAGGCCGAAGTGCAGTTGGTTGAATCAGGTGGGGGTT TAGTACAGCCGGGTGGTAGTTTACGTTTGTCATGTGCGGCATCAGGTTTTATTTTTAGTAGTGATTGGATGAATTGGGTA CGTCAAGCACCGGGAAAAGGATTAGAATGGGTGGCGAATATTAATCAAGATGGTTCAGAAAAATATTATGTGGATTCAGT

TAAAGGTCGTTTTACAATCAGCCGTGACAACGCACAAAATAGCTTATACTTACAAATGAACAGTTTACGGGCAGAAGACA
CAGCAGTATATTATTGTGCAAAGGAATTAGGGCCGTGGGGGCAAGGGACATTAGTGACGGTGAGTAGCGCCTCCACCAAG
GGCCCATCGGTCTTCCCCCTGGCACCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGA
CTACTTCCCCGAACCGGTGACGGTGTCGTGGAACTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCCTAC
AGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACTGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAAC
GTTAATCACAAACCCAGCAACACCAAGGTCGACAAGAAAGTTGAGCCCAAATCTTGCCATCACCACCATCACCATTAATA
ACCATGGAGAAAATAAAGTGAAACAAAGCACTATTGCACTGGCACTCTTACCGTTACTGTTTACCCCTGTGACAAAAGCC
caggcagtagttattcaggaaagcgcACTCACGACATCTCCGGGGGGGACGGTTATTCTCACTTGCCGCAGCAGTACAGG
AACGATTACGACTTCTAACTATGCAAATTGGGTCCAGAAAAAACCGAATCATGTGTTTACGGGTTTAATTGGGGCAACGA
GCATTCGCGCGCCGGGAGTGCCGGTACGTTTTAGCGGGTTTCTTATTGGTGGAAAGGCAGCATTAACTATTACAGGAGCG
CAAACCGAAGATGATGCTATGTATTTTTGCGCGTTATGGTATAACACACACTATGTTTTTGGAGGTGGCACGAAGGTTAC
AGTATTGGGCCAGCCGAAAGCCGccCCCTCCGTCACCCTGTTCCCTCCCTCCTCTGAAGAACTGCAGGCTAACAAAGCGA
CCCTGGTTTGTTTAATTTCTGACTTCTACCCTGGTGCCGTCACCGTAGCATGGAAAGCCGATTCATCCCCCGTAAAAGCC
GGAGTGGAGACCACGACCCCCTCAAAACAGTCCAATAACAAATATGCCGCCTCTTCCTATCTCAGCCTTACCCCCGAACA
ATGGAAATCACACCGTTCTTATTCCTGCCAAGTCACCCACGAAGGTTCCACCGTTGAAAAAACCGTAGCCCCAACCGAAT
GCTCCTAATAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCGTTTACCGCTACT
GCGTCACGGATCTCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGC
TCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATG
GTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTC
TTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTA
TTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCA
CTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAA
TAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTT
TTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG
CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATG
ATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCAT
ACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT
TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTA
ACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA
CGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG
CTTCCCGGCAACAATTGATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGC
TGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGCTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCC
CTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG
CCTCACTGATTAAGCATTGGTAGGAATTAATGATGTCTCGTTTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTG
CTTAATGAGGTCGGAATCGAAGGTTTAACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATTG
GCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTAGATAGGCACCATACTCACTTTTGCCCTT
TAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAACGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGA
GCAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAGTATGAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACA
AGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCAGTGGGGCATTTTACTTTAGGTTGCGTATTGGAAGATCAAG
AGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCGCCATTATTACGACAAGCTATCGAATTA
TTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGAAAAACAACTTAAATG
TGAAAGTGGGTCTTAAAAGCAGCATAACCTTTTTCCGTGATGGTAACTTCACTAGTTTAAAAGGATCTAGGTGAAGATCC
TTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAA
GGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTG
TTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGC
TATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC
ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATT
TTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCT
GGCCTTTTGCTCACATG

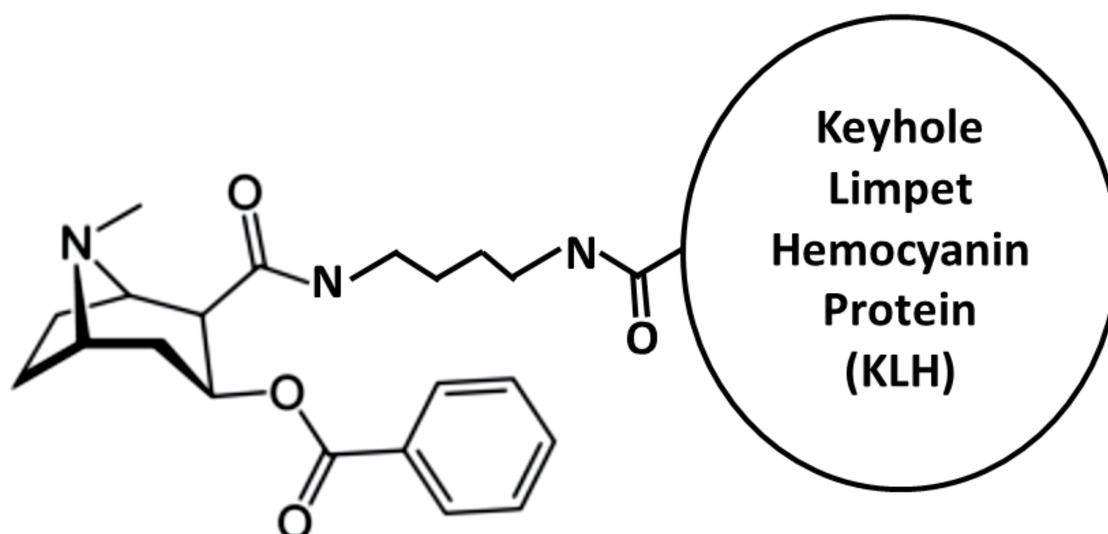
Table S5 Elbow angles observed in crystals of Fab-BE complex and unliganded Fab.

rFab complex crystals	
Fab (light, heavy chain id)	Elbow angle (°)
Fab-1 (L, H)	151.8
Fab-2 (A, B)	151.9
Fab-3 (C,D)	155.8
Fab-4 (E,F)	165.6
Fab-5 (I, J)	173.1
Fab-6 (M, N)	166.1
Fab-7 (O, P)	171.6
Fab-8 (J, K)	163.4
Free mFab crystals	
Fab (light, heavy chain id)	Elbow angle (°)
Fab-1 (L, H)	145.1
Fab-2 (A, B)	137.1

Table S6 The ligand binding site is occluded for one of the two Fab molecules found in the asymmetric unit of the apo Fab crystals. This table lists the hydrogen bonds observed between one of the Fab molecules in the asymmetric unit (chain A, light chain) and atoms from a crystallographic contact (chain B, heavy chain). In addition to the hydrogen bonds, aromatic – pi interaction was observed between Trp93A and Lys129B of the crystallographic contact.

Fab atoms from the asymmetric unit	Atoms from crystal contact	Distance (Å)
Tyr34A OH	Ser130B N	2.6
Trp93A NE1	Ser130B O	3.1

Figure S1 Benzoylecgonine (BE)-KLH antigen used to generate the anti-cocaine h2E2 mAb.



S2. CHO-cell production of the recombinant mAb h2E2

The hybridoma cell line secreting the original anti-cocaine mAb designated 2E2 that was composed of a human γ1 heavy chain and a murine λ light chain was generated, identified and isolated using standard hybridoma technology. This was accomplished as previously described by fusing splenocytes obtained from a transgenic mouse strain, designated HCo7/Ko5 that had been immunized with a hapten-carrier conjugate with cells of the murine myeloma cell line P3X63-Ag8.653 (Paula *et al*., 2004; Fishwild *et al*., 1996). Further, the re-engineering and replacement of the murine λc light chain constant region with the human λc domain and use of Catalent Pharma (Madison, WI) GPEx® technology to generate stably transfected CHO-S cells that produce high levels of the recombinant, further "humanized" h2E2 mAb as well as its purification has been previously described (Norman *et al*., 2014).

S3. Preparation and characterization of Fab from mAB h2E2 and reductive methylation of Fab fragment

The h2E2 anti-cocaine mAb Fab fragment was generated by Endo-Lys-C digestion, followed by purification, as described previously (Kirley & Norman 2015). The Fab fragment was reductively methylated on lysine residues as a way to aid crystallization, basically as described previously (Walter *et al*., 2006; Tan *et al*., 2014). Typically, 10 mg of purified Fab was cooled to 4°C in 10 ml of 50 mM HEPES, 250 mM NaCl, pH=7.5. To that sample, 0.2 ml of 1.0M Borane-Dimethyl complex ("ABC" complex, in water) and 0.4 ml of 1.0M formaldehyde (in water) were added, the solution gently mixed, and incubated for 2 hours at 4°C. Then, an additional 0.2 ml of 1.0M Borane-Dimethyl complex ("ABC") in water and 0.4 ml of 1.0M formaldehyde in water were added, the solution gently mixed, and the incubation continued at 4°C for 2 more hours. Finally, an additional 0.1 ml 1.0M Borane-Dimethyl complex ("ABC") was added, the solution gently mixed, and the reaction allowed to continue overnight at 4°C (an additional 18 hours). The approximately 10 ml reaction mixture was concentrated to approximately 2 ml using an Amicon Ultra 15ml 30 kDa MWCO spin concentrator, and immediately loaded onto a 43 ml, (28 cm long) Sephacryl S-100 size exclusion chromatography column equilibrated in 20 mM Tris-Cl, 100 mM NaCl, pH=7.2. The column was run at ambient temperature at 0.5 ml/min and fractions collected every 2 minutes. Fractions were examined for Abs280nm, and the peak fractions (25-30) were pooled and concentrated to less than 1 ml using an Amicon Ultra 15ml 30 kDa MWCO spin concentrator. The concentrated methyl-Fab preparation was quantified by dilution and measurement of A280nm, using the extinction coefficient for the purified Fab of 73,965 $M^{-1}cm^{-1}$ (calculated from the amino acid sequence, and equivalent to OD 280nm = 1.612 for a 1.0 mg/ml solution of Fab).

Fab and methylated Fab samples were buffer exchanged into 0.1% formic acid immediately before submission for mass spectral analyses, which were done in the laboratory of Dr. Ken Greis at the University of Cincinnati, College of Medicine, on a fee for service basis. Flow infusion electrospray ionization time-of-flight (ESI-TOF) mass spectrometry was employed to determine protein masses. Briefly, each sample, at approximately 2-4 mg/ml, was diluted 1:1 in 60% acetonitrile/0.1% formic acid to a final concentration of 30% ACN/0.1% formic acid. The sample was infused into the mass spectral ABSciex 5600 plus system at 500nL/min and the positive ion Tof-MS spectra were collected. The data were further processed using the protein reconstruct algorithm in the PeakView ver 2.1 software from ABSciex to generate the reconstructed mass profile. Apomyoglobin (mass = 16951.3) was also infused at 2.5 pmole/uL in the same solvent to validate the process and validate the reconstructed mass.

Both Fab and methylated Fab (mFab) were examined by ESI-TOF mass spectral analysis. Based on the amino acid sequence, the presence of 5 disulfide bonds per Fab fragment, and a glutamine to pyroglutamate post-translational modification of the N-terminus of the Fab light chain, the theoretical calculated average mass of the underivatized h2E2 Fab fragment is 45865.5 Da. Four independent Fab preps yielded experimental masses of 45867.8, 45867.2, 45866.9, and 45468.8 Da (average = 45867.7 ± 0.9 Da), possibly suggesting two amidations on the Fab fragment (delta mass per amidation = 1.0 Da). Three preparations of methyl-Fab were produced for mass

spectral analyses and crystallization trials, and the purified yield of the methyl-Fab was 81% ± 7% of the starting Fab. To determine the number of methyl groups added by the methylation reactions, paired starting Fab fragments and methylated Fab fragments were analyzed. For three methyl-Fab preparations, the delta mass observed between the methyl-Fab and the Fab fragment were +686, +689, and +689 Da. Since each methyl group adds +14 Da, this corresponds to 49.0, 49.2, and 49.2 methyl groups added per Fab. The maximum number of methyl groups that theoretically could be introduced by methylation of the Fab fragment = 2 X (24 lysine primary amines + 1 heavy chain alpha amino primary amine) = 50 possible methyl groups (each primary amine can incorporate 2 methyl groups, and the light chain alpha amino primary amine is not present due to the pyroglutamate post-translational modification of the N-terminus of the light chain.) Thus, methylation is very close to complete, with 49/50 possible methyl groups added to the Fab fragment by the procedure. This is similar to what has been reported earlier by others (Walter *et al*., 2006).